2/27 8

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
JAMES DREVER (Nos. 1–2)
BORIS SEMEONFF (Nos. 3–4)

WITH THE ASSISTANCE OF
C. BURT

AIDED IN THE SELECTION OF PAPERS BY

F. C. BARTLETT
E. A. BOTT
R. S. CREED
G. C. DREW
L. S. HEARNSHAW
M. G. KENDALL
T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS

ASSISTANT EDITOR: IAN M. L. HUNTER

VOLUME 49, 1958

CAMBRIDGE
AT THE UNIVERSITY PRESS
1958

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York 22, N.Y.
Agent in Canada: The Macmillan Company of Canada Limited

Bureau Edn. Research
DAVID ...
Dated 27. 8. 70 .....
Acc. No. J 231 .....

CONTENTS OF VOL. 49

PART 1. FEBRUARY 1958

PAGE

A factorial study of introversion-extraversion. By H. P. HILDEBRAND . 1

Determinants of emotionality in the rat. II. Antecedent factors. By P. L. BROADHURST 12

Surprise, isolation and structural change as factors affecting recall of a temporal series. By ROBERT T. GREEN 21

Impossible objects: a special type of visual illusion. By L. S. PENROSE and R. PENROSE 31

The effects of the menstrual cycle on test performance. By MARY WICKHAM 34

A 'Crespi effect' in male imbeciles. By N. O'CONNOR and G. S. CLARIDGE . 42

The effects of incentives on the performance of defective imbeciles. By D. WALTON and T. L. BEGG 49

A 'Crespi effect' in the analysis of emotionality as a drive in rats. By P. L. BROADHURST 56

A Pattern Recognition method of learning morse code. By M. D. ALLAN . 59

The development of independent muscular control of separate eyelids in two racial groups, European and Polynesian. By P. R. N. SUTTON . . 65

Shape recognition: a reply to Dodwell. By J. A. DEUTSCH 70

Publications recently received 72

Other publications received 85

PART 2. MAY 1958

Partial advance information in a choice reaction task. By J. ALFRED LEONARD 89

Factors affecting inductive predictions. By R. T. GREEN 97

Experimental design: a cloak for intellectual sterility. By J. G. TAYLOR . 106

The effect of set on accuracy of auditory perception. By GEORGE A. TALLAND 117

Tachistoscopic exposure time and spatial proximity in the organization of visual perception. By J. W. KASWAN 131

Psychomotor control in stammering and normal subjects: an experimental study. By GEORGE SETH 139

Cognitive and social changes in the feeble-minded—three further studies. By A. D. B. CLARKE, A. M. CLARKE and S. REIMAN 144

Shape recognition: a reply to Deutsch. By P. C. DODWELL 158

Critical notice. By JEROME S. BRUNER 160

Publications recently received 164

Part 3. August 1958

PAGE

Obituary Notice: Ernest Jones 177

Visually controlled locomotion and visual orientation in animals. By James J. Gibson 182

Experiments on 'neophobia' in wild and laboratory rats. By S. A. Barnett 195

The role of the reinforcer as a stimulus. By R. L. Reid 202

Visual search for successive decisions. By N. H. Mackworth and J. F. Mackworth 210

The role of attitudes in comprehension. By S. Friedmann 222

Time for reading and memory. By E. C. Poulton 230

Effects of practice and rest on fluctuations in the Müller–Lyer illusion. By H. J. Eysenck and P. Slater 246

Publications recently received 257

Other publications received 264

Part 4. November 1958

Binocular resolution and perception of race in South Africa. By Thomas F. Pettigrew, Gordon W. Allport and Eric O. Barnett . . . 265

Attention to visual displays during a vigilance task. I. Biasing attention. By C. H. Baker 279

Exploratory behaviour. By S. A. Barnett 289

The attention-getting value of structural change. By Robert T. Green . 311

The influence of the albedo and complexity of stimuli on visual fixation in the human infant. By D. E. Berlyne 315

Skill and chance: variations in estimates of skill with an increasing element of chance. By John Cohen, E. J. Dearnaley and C. E. M. Hansel . 319

Effect of quinal barbitone, dextro-amphetamine, and placebo on apparent time. By Sanford Goldstone, William K. Boardman and William T. Lhamon 324

Extreme response sets as a measure of intolerance of ambiguity. By M. I. Soueif 329

Drive strength and adaptation to stress. By H. M. B. Hurwitz and J. Rowell 335

The meaning of 'intelligence' to Lebanese teachers. By Jack D. Keehn and E. Terry Prothero 339

Publications recently received 343

INDEX OF AUTHORS

	PAGE
Allan, M. D.	59
Allport, G. W.	265
Baker, C. H.	279
Barnett, E. O.	265
Barnett, S. A.	195, 289
Begg, T. L.	49
Berlyne, D. E.	315
Boardman, W. K.	324
Broadhurst, P. L.	12, 56
Bruner, J. S.	160
Claridge, G. S.	42
Clarke, A. D. B.	144
Clarke, A. M.	144
Cohen, J.	319
Dearnaley, E. J.	319
Deutsch, J. A.	70
Dodwell, P. C.	158
Eysenck, H. J.	246
Friedmann, S.	222
Gibson, J. J.	182
Glover, E.	177
Goldstone, S.	324
Green, R. T.	21, 97, 311
Hansel, C. E. M.	319
Hildebrand, H. P.	1
Hurwitz, H. M. B.	335
Kaswan, J. W.	131
Keehn, J. D.	339
Leonard, J. A.	89
Lhamon, W. T.	324
Mackworth, J. F.	210
Mackworth, N. H.	210
O'Connor, N.	42
Penrose, L. S.	31
Penrose, R.	31
Pettigrew, T. F.	265
Poulton, E. C.	230
Prothro, E. T.	339
Reid, R. L.	202
Reiman, S.	144
Rowell, J.	335
Seth, G.	139
Slater, P.	246
Soueif, M. I.	329
Sutton, P. R. N.	65
Talland, G. A.	117
Taylor, J. G.	106
Walton, D.	49
Wickham, M.	34

INDEX OF PUBLICATIONS REVIEWED

Abell, W. *The Collective Dream in Art* 174
Alexander, F. *Psychoanalysis and Psychotherapy* 77
Anderson, C. C. *Function Fluctuation* 349
Ansbacher, H. L. & R. R. *The Individual Psychology of Alfred Adler* 348
Argyle, M. *The Scientific Study of Social Behaviour* 74
Bartlett, F. C. *Some Recent Developments of Psychology in Great Britain* 80
Bartlett, F. C. *Thinking* 160
Bartley, S. H. *Principles of Perception* 257
Becker, H. & Boskoff, A. *Modern Sociological Theory* 81
Berg, C. & Krich, A. M. *Homosexuality* 261
Boss, M. *The Analysis of Dreams* 83
Bott, E. *Family and Social Network* 171
Bottome, P. *Alfred Adler* 83
Brown, C. C. & Saucer, R. T. *Electronic Instrumentation for the Behavioral Sciences* 349
Bruner, J. S. and others. *Contemporary Approaches to Cognition* 257
Cattell, R. B. *Personality and Motivation Structure and Measurement* 350
Cleugh, M. F. *The Slow Learner* 167
Clifford, A. *Homosexuality* 261
Connolly, T. G. & Sluckin, W. *Statistics for the Social Sciences* 175
Davis, D. R. *An Introduction to Psychopathology* 169
Dews, P. B. and others. *Techniques for the Study of Behavioral Effects of Drugs* 84
Diamond, S. *Personality and Temperament* 261
Ergonomics, Vol. 1, no. 1 175
Eysenck, H. J. *Sense and Nonsense in Psychology* 82
Eysenck, H. J. *The Dynamics of Anxiety and Hysteria* 162
Festinger, L. *A Theory of Cognitive Dissonance* 258
Fletcher, R. *Instinct in Man* 174
Foulkes, S. H. & Anthony, E. J. *Group Psychotherapy* 170
Frandsen, A. N. *How Children Learn* 82
Freeman, T., Cameron, J. L. & McGhie, A. *Chronic Schizophrenia* 262
Fromm, E. *The Sane Society* 76
Garfield, S. L. *Introductory Clinical Psychology* 78
Greenblatt, M., Levinson, D. J. & Williams, R. H. *The Patient and the Mental Hospital* 263
Hamlyn, D. W. *The Psychology of Perception* 73
Hunter, I. M. L. *Memory* 74
International Sociological Association. *The Nature of Conflict* 172
Jung, C. G. *Symbols of Transformation* 83
Kety, S. S. and others. *The Pharmacology of Psychotomimetic and Psychotherapeutic Drugs* 84
Kline, N. S. *Psychopharmacology* 84

KRECH, D. & CRUTCHFIELD, R. S. *Elements of Psychology* 259
LONERGAN, B. J. F. *Insight* 173
LOVELL, K. *Educational Psychology and Children* 344
MACINTYRE, A. C. *The Unconscious* 260
MAXWELL, A. E. *Experimental Design in Psychology and the Medical Sciences* 259
MUNROE, R. L. *Schools of Psychoanalytic Thought* 164
O'NEIL, W. M. *Introduction to Method in Psychology* 259
OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. *The Measurement of Meaning* 72
PEAR, T. H. *Personality, Appearance and Speech* 75
PIOTROWSKI, Z. A. *Perceptanalysis* 172
PODOLSKY, E. *The Neuroses and their Treatment* 347
REEVES, J. W. *Body and Mind in Western Thought* 343
SHUEY, A. M. *The Testing of Negro Intelligence* 347
STEINFELD, J. I. *A New Approach to Schizophrenia* 345
STENGEL, E. & COOK, N. G. *Attempted Suicide* 346
VERNON, M. D. *Backwardness in Reading* 166
WALKER, N. *A Short History of Psychotherapy* 76
WALSH, E. G. *Physiology of the Nervous System* 78
WHITAKER, C. *Psychotherapy of Chronic Schizophrenic Patients* 345
WILNER, D. M., WALKLEY, R. P. & COOK, S. W. *Human Relations in Interracial Housing* 81
WINNICOTT, D. W. *The Child and the Family* 83
WINNICOTT, D. W. *The Child and the Outside World* 83
WOLSTENHOLME, G. E. W. *Colloquia on Ageing, Vol.* 3 80
WOLSTENHOLME, G. E. W. & MILLER, E. C. P. *Extrasensory Perception* 79
ZAEHNER, R. C. *Mysticism, Sacred and Profane* 83

A FACTORIAL STUDY OF INTROVERSION–EXTRAVERSION

BY H. P. HILDEBRAND
Tavistock Clinic

ABSTRACT

Previous research suggested that Jung's Theory of Psychological Types could best be examined by factorial methods. Two orthogonal factors of neuroticism and introversion-extraversion were hypothesized in order to explain Jung's theory. A varied neurotic population was examined on a large battery of tests, and factors identified as 'neuroticism', 'g' and 'introversion-extraversion' extracted. Further evidence confirming Jung's theories was given by using regression equations and factor scores which enabled criterion groups not included in the original matrix to be correctly identified.

I. INTRODUCTION: DEFINITION

Jung (1913) is commonly acknowledged to have coined the terms 'Introversion' and 'Extraversion' to describe a typology which he traces back to its earliest beginnings—the well-known division into 'habitus phythisicus' and 'habitus apoplecticus' elaborated by Galen. Other modern authors before Jung, notably Jordan (1890) and Gross (1902), had put forward psychological theories based on a type division, but Jung was the first to gain widespread acceptance for the type theory. Since his descriptions of the introvert and the extravert have been the basis for much of the subsequent work in this field, it would seem advisable to quote them in some detail.

Introversion: means a turning inward of the libido whereby a negative relation of subject to object is expressed. Interest does not move towards the object but recedes toward the subject. Everyone whose attitude is introverted thinks, feels and acts in a way that clearly demonstrates that the subject is the chief factor of motivation while the object at most receives only a secondary value. (Introversion may possess either a more intellectual or a more emotional character) just as it can be characterized by either intuition or sensation. Introversion is active when the subject *wills* a certain seclusion in the face of the object: it is passive when the subject is unable to restore again to the object the libido which is streaming back from it. When Introversion is habitual, one speaks of an *introverted type* (1924, p. 567).

Extraversion: means an outward turning of the libido. With this concept I denote a manifest relatedness of subject to object in the sense of a positive movement of subjective interest towards the object. Everyone in a state of extraversion thinks, feels and acts in relation to the object, and moreover in a direct and clearly observable fashion, so that no doubt can exist about his positive dependence upon the object. In a sense, therefore, extraversion is an outgoing transference of interest from the subject to the object. If it is an intellectual extraversion, the subject thinks himself into the object; if a feeling extraversion then the subject feels himself into the object. The state of extraversion means a strong, if not exclusive, determination by the object. One should speak of an *active* extraversion when deliberately willed, and of a *passive* extraversion when the object compels it, i.e. attracts the interest of the subject of its own accord, even against the latter's intention. Should the state of extraversion become habitual the *extraverted* type appears (1924, p. 543).

Three further points are of importance. First, Jung states that

although there are individuals in whom one can recognize the type at first glance, this is by no means always the case. As a rule only careful observation and weighing of the evidence permits a sure classification. However clear and simple the fundamental principle of the opposing attitudes may be, this concrete reality is none the less complicated and obscure, for every individual is an exception to the rule. Therefore, one can never give a description of a type, no matter how complete, which absolutely applies to one individual, despite the fact that thousands might, in a sense, be characterised by it.

A FACTORIAL STUDY OF INTROVERSION–EXTRAVERSION

BY H. P. HILDEBRAND
Tavistock Clinic

ABSTRACT

Previous research suggested that Jung's Theory of Psychological Types could best be examined by factorial methods. Two orthogonal factors of neuroticism and introversion-extraversion were hypothesized in order to explain Jung's theory. A varied neurotic population was examined on a large battery of tests, and factors identified as 'neuroticism', 'g' and 'introversion-extraversion' extracted. Further evidence confirming Jung's theories was given by using regression equations and factor scores which enabled criterion groups not included in the original matrix to be correctly identified.

I. INTRODUCTION: DEFINITION

Jung (1913) is commonly acknowledged to have coined the terms 'Introversion' and 'Extraversion' to describe a typology which he traces back to its earliest beginnings—the well-known division into 'habitus phythisicus' and 'habitus apoplecticus' elaborated by Galen. Other modern authors before Jung, notably Jordan (1890) and Gross (1902), had put forward psychological theories based on a type division, but Jung was the first to gain widespread acceptance for the type theory. Since his descriptions of the introvert and the extravert have been the basis for much of the subsequent work in this field, it would seem advisable to quote them in some detail.

Introversion: means a turning inward of the libido whereby a negative relation of subject to object is expressed. Interest does not move towards the object but recedes toward the subject. Everyone whose attitude is introverted thinks, feels and acts in a way that clearly demonstrates that the subject is the chief factor of motivation while the object at most receives only a secondary value. (Introversion may possess either a more intellectual or a more emotional character) just as it can be characterized by either intuition or sensation. Introversion is active when the subject *wills* a certain seclusion in the face of the object: it is passive when the subject is unable to restore again to the object the libido which is streaming back from it. When Introversion is habitual, one speaks of an *introverted type* (1924, p. 567).

Extraversion: means an outward turning of the libido. With this concept I denote a manifest relatedness of subject to object in the sense of a positive movement of subjective interest towards the object. Everyone in a state of extraversion thinks, feels and acts in relation to the object, and moreover in a direct and clearly observable fashion, so that no doubt can exist about his positive dependence upon the object. In a sense, therefore, extraversion is an outgoing transference of interest from the subject to the object. If it is an intellectual extraversion, the subject thinks himself into the object; if a feeling extraversion then the subject feels himself into the object. The state of extraversion means a strong, if not exclusive, determination by the object. One should speak of an *active* extraversion when deliberately willed, and of a *passive* extraversion when the object compels it, i.e. attracts the interest of the subject of its own accord, even against the latter's intention. Should the state of extraversion become habitual the *extraverted* type appears (1924, p. 543).

Three further points are of importance. First, Jung states that

although there are individuals in whom one can recognize the type at first glance, this is by no means always the case. As a rule only careful observation and weighing of the evidence permits a sure classification. However clear and simple the fundamental principle of the opposing attitudes may be, this concrete reality is none the less complicated and obscure, for every individual is an exception to the rule. Therefore, one can never give a description of a type, no matter how complete, which absolutely applies to one individual, despite the fact that thousands might, in a sense, be characterised by it.

Secondly, following the implications that I–E is a continuum, Jung identifies the poles of the distribution of I–E with two of the main groups of neurotic disorders—the extravert with the conversion hysterias, and the introvert with Janet's psychasthenias. He says (1924, p. 421); 'Hysteria is in my view, by far the most frequent neurosis with the extraverted type', while psychasthenia is identified as the typical neurotic disorder of the introvert. This neurosis is characterized on the one hand by marked sensitivity, and on the other by great exhaustion and constant tiredness. Eysenck has suggested the term 'dysthymic' to cover this syndrome of affective disorders. In modern psychiatric terminology dysthymic would cover such syndromes as anxiety states and reactive depressions.

Thirdly, Jung stresses the independence of introversion and neurosis: he says—'It is a mistake to believe that introversion is more or less the same as neurosis. As concepts the two have not the slightest connexion with one another.'

Jung's formulation was widely accepted and a great number of studies carried out in an attempt to provide an experimental basis for the theory. But as Collier & Emch (1938) point out, there was considerable confusion as to the definition of the traits it was hoped to define. The main lines of Jung's arguments were ignored, while the terms introversion and extraversion were retained and redefined according to the authors' theoretical preferences. Collier and Emch remark that there are three main varieties of redefinitions which still claim to have some connexion with Jung's original dichotomy. First, there is the attempt at simplification; secondly, the identification of introversion with neurotic tendency—extraversion being given 'per contra' a connotation of normality; and thirdly, the relating of introversion to the schizophrenic and extraversion to the manic depressive psychoses.

Not surprisingly the experimental evidence is very inconclusive. Experiments based on the first variety of redefinitions are quite inadequate to explain the wide variety of phenomena cited by Jung. The second type of definition is based on Freud's adaptation of the term introversion. For Freud, introversion is 'one of the invariable and indispensable considerations in every case of psycho-neurosis—a substitution for actual objects of phantasies of these objects' (Freud, 1920). It is Freud's belief that Jung's use of the term is 'indiscriminate'. As Collier and Emch point out 'Freud's definition of the term represents a rigid contraction both in meaning and in use'. They maintain that Freud's definition of introversion has been particularly pervasive in the work of writers interested in the subject, but paradoxically the writers concerned seem to regard it as a natural extension of Jung's concept rather than a contraction. It is beyond the scope of this paper to examine this aspect of the controversy, but it is apparent that any work based on Freud's use of the term introversion can have little relation to the original formulation. Similarly, of course, any attempt to identify introversion with psychotic manifestations goes far beyond Jung's formulations, and can be disregarded in terms of validation of the hypothesis.

There seems no point in carrying the argument further, particularly as all early experimental work is summed in three excellent reviews of work carried out to establish the empirical basis of I–E. They are those of Guilford (1934), Collier & Emch (1938), and Eysenck (1947), and are devoted in the main to small-scale experiments. Two further causes of confusion presented by these authors are, first, the clinical identification of introversion with neurosis referred to above, and secondly, the identification of extra-

version with sociability first put forward by Freyd (1924). This identification was disproved when put to an empirical test by Russell Frazer (1947). Lack of sociability must be regarded as an index of neuroticism, not as a sign of introversion (Eysenck, 1947).

Eysenck (1953*a*) has also presented a review of factorial studies with some bearing on Jung's hypothesis. In the main these studies agreed in finding 'a bi-polar factor variously called "surgency-desurgency" (Cattell, 1946), "c" (Garrett, 1921), "aggressive-inhibitive" (Line, 1935) or whatever terms seem appropriate...these factors closely agree with one another as well as our own hysteria-dysthymia factor'. All the studies named above are based upon rating scales and have their origin in the pioneer work of Heymans & Wiersma (1906) and of Webb (1915). The name given to a factor arising from the intercorrelation of a number of traits seemed mainly to depend on the theoretical bias of the author. The low reliability and validity of such studies make caution necessary when identifying such factors with Jung's Introversion-Extraversion hypothesis, but the similarity of the results is striking.

Further evidence is available from questionnaire studies. Of these the questionnaire constructed by Guilford is outstanding in terms of reliability and soundness of statistical construction. Guilford himself would explain I–E in terms of a number of oblique factors, but Lovell (1945) and North (1948) have shown that the questionnaire in fact gives rise to two second-order factors, the second of which shows a distinct resemblance to the factors found by the rating method (Hildebrand, 1956).

II. THE PRESENT STUDY

(i) *Rationale*

In the brief summary of previous work reported above, it has been shown that while small-scale testing has generally been ineffective to test Jung's hypotheses, various factorial studies have reported factors which seemed to resemble I–E. The evidence in favour of such a factor is suggestive but not convincing. In the present study, however, it was decided to test Jung's hypothesis directly, using factorial methods and a battery of tests designed to describe introversion-extraversion as completely as possible. The following hypotheses were derived from Jung's original statement by the writer. *Hypothesis* 1: That I–E (Introversion-Extraversion) is an underlying dimension of personality, common in some degree to all men. *Hypothesis* 2: That I-E is distinct from a second underlying dimension, namely, Neuroticism. *Deduction* 2. 1: That if a factor analysis is carried out on a group of tests purporting to measure I-E and Neuroticism two general factors should emerge, which can most meaningfully be explained as I-E and Neuroticism. *Deduction* 2. 2: That these two factors would be significant and orthogonal. *Hypothesis* 3: That extraverts would tend as a group to suffer from a hysterical type of neurosis; that introverts as a group would tend to suffer from an anxiety neurosis. *Deduction* 3. 1: That hysterics as a group would tend to score high on the Neuroticism factor and also tend to cluster at the E pole of the I-E factor; conversely although anxiety states as a group would also tend to score high on the Neuroticism factor they would tend to cluster at the I pole of the I-E factor. *Deduction* 3. 2: That mixed neurotic groups would score high on the Neuroticism factor, but would tend towards a central distribution on the I-E factor. *Deduction* 3. 3: That normal subjects would have a central tendency on the I-E factor and low scores on the Neuroticism factor. The deductions made from these hypotheses are available to proof or disproof by experimental investigation.

In view of the nature of the investigation it was decided to use so-called objective behaviour tests. ('Objective behaviour tests', as here used, implies only that the tests are scored objectively and not that the tests themselves are objective. Such tests are scored by a standard set of instructions which do not require subjective interpretation by the interpreter.) The reasons for this decision were that the use of such tests would enable the experiment to be repeated independently if required, that the reliability and validity of the tests were usually known and that no evidence was at that time available about the interaction of these tests with regard to the factors under investigation.

According to the criteria laid down above, two groups of tests were chosen, one of which had been shown by previous experiments to differentiate between groups of normals and neurotics, and a second which had been shown to differentiate significantly between dysthymics and hysterics in previous investigations. It was, of course, necessary that such differentiation did not depend upon subjective judgments but was a function of the test scores themselves. A group of tests with high loadings on the third factor 'g' was also included as previous work had suggested that there was some relationship between I-E and intelligence (Eysenck, 1947).

It was proposed to resolve the crucial point of the generality of the I-E factor in the following manner. If the writer's interpretation of Jung's hypothesis was correct, then the exclusion of the two groups, which he describes as being the most typical of the poles of the dimension, from a factorial experiment should not affect the extraction of a factor which would be most meaningful psychologically when described as I-E. The hysteric and anxiety groups would be treated as criterion groups and not included in the factor analysis. Similarly, the normal control group was regarded as a criterion group for the neuroticism factor.

After the analysis had been carried out it was further proposed to make estimations of the regression coefficients of each variable, and to plot the groups in a two-factor space with regard to the two factors of Neuroticism and I-E. If the anxiety states and hysterics were found to be significantly different on the I-E factor after their score had been estimated and not significantly different on the Neuroticism factor then deduction 3. 1 would be proved. A plot of the mixed neurotic groups would afford proof or disproof of Deduction 3. 2 and a plot of the normals would do the same for Deduction 3. 3.

An additional strength of this method is that it would confirm or disprove differentiation between criterion groups by means of analysis of variance techniques, and indicate where such differentiation was due to specific causes or the action of the I-E factor. It was also decided to include subjects suffering from all types of neurotic disorder, in order that the different syndromes could be related to the relevant personality dimensions of Neuroticism and I-E. A normal control group was also tested. It should be noted that the results found are those from a neurotic population and cannot immediately be generalized to cover the behaviour of a normal population.

(ii) *Subjects*

All neurotic subjects were drawn from in-patients at Belmont Hospital, Sutton, Surrey. All subjects were male. Minski gives the following details (1950): 'Approximately 1500 to 2000 patients are admitted to Belmont annually for an average length of stay of 2 to 3 months. By far the majority of admissions are neurotics and come from all over the country. All patients are voluntary.'

All admissions over an 8-month period were screened. From these 145 patients were selected as falling within the following categories. (1) Over 18 and under 50 years of age. (2) Considered by psychiatrist

to show no sign of organic lesion. (3) Considered by psychiatrist to show no sign of psychotic illness. (4) I.Q. above 90. (5) Not undergoing physical treatment (E.C.T. or Insulin Therapy). All patients were severe neurotics, having been given up by both G.P. and out-patient clinics, and the amount of mixed symptomatology was great. Cases where the psychiatrist felt justified in making a clear-cut diagnosis were few, and the groups into which the patients were sorted should be regarded as classifications in terms of outstanding syndromes.

Details of the breakdown of subjects into groups are given in Table 1. The total number of patients tested was 145. Overall analyses of variance between the groups showed no significant difference in age. As a group, however, the Obsessionals and Depressives tended to be somewhat older. The mean age was 33·9 years, with a standard deviation of $\pm 8{\cdot}1$.

Table 1

Group	No.	Abbreviation
Conversion hysterics	25	Dys.
Psychopaths	20	P.P.
Reactive depression	10	Dep.
Obsessional neurosis	10	Obs.
Anxiety states	25	Anx.
Mixed symptomatology	55	Mx.

The normal control group consisted of 25 male soldiers taken at random from a reallocation centre in the neighbourhood. Eysenck states of a preceding and similar group to that tested by the writer: 'It is known that soldiers going through such a centre tend to have a higher degree of neurotic and psychotic abnormality than do soldiers not reallocated. In addition these soldiers tend to be slightly below average in intelligence, and therefore constitute a distinctly poor sample of the normal population.' That it is possible, in spite of this difficulty, to obtain highly significant differentiation between such a group and abnormal groups has been shown already (Eysenck, 1953*b*). The mean age of the Normal group equals $24{\cdot}5 \pm 6{\cdot}6$ years. The C.R. between the mean age of this group and the mean age of the neurotic group is significant (C.R. = 5·53). However, the difficulty of obtaining a homogeneous normal sample imposed the use of this group. The testing of the control group was carried out at the Maudsley Hospital, and the same procedure was used as with the neurotic groups at Belmont.

(iii) *Experimental variables*

In order to test the hypotheses propounded above a review was made of the literature and three groups of tests selected. The first of these was a group of three intelligence tests. Eysenck had suggested that intelligence might have some bearing on I-E and this group of tests was selected as covering the cognitive field. The tests were the Progressive Matrices, the Mill Hill Vocabulary and the Nufferno Power Test. Detailed descriptions of all tests and the rationale for use may be found in the writer's Ph.D. thesis (Hildebrand, 1953).

The second group of tests consisted of tests which it was hoped would define the Neuroticism factor. They were as follows: (4) The Maudsley Medical Questionnaire. (5) The Maudsley Medical Questionnaire —lie scale. (6) An annoyances Questionnaire—a list of thirty possible annoyances adapted from Pressey's X–O test. Previous research suggested that neurotics would be more irritable than normal groups.

The third group of tests included those which previous workers had suggested might have some bearing on the I-E dimension. They were (7) A Humour Scale—subjects were presented with thirty cartoons and asked to classify them according to preference. There were ten cartoons each for Sex humour, Incongruous humour, and Extra-punitive humour. (8) The Guilford Martin Inventory of Factors S T D C R. Factors S and R had been found by previous workers to have some bearing on I–E, and factors T D and C on neuroticism (Hildebrand, 1956). (9) Static ataxia and (10) Suggestibility: four scores each were taken for static ataxia and for suggestibility to autosuggestion—maximum forward sway, maximum backward sway; total sway in half-inch units; and number of reversals (oscillations) in quarter-inch units. (11) The Porteus maze Test—adapted Q score. Scoring categories were Starting time, Tracing time, Weighted wrong directions, Lifted pencils, Crossed lines. (12) Ratio stature/Transverse chest diameter in inches. (13) Leg persistence—the time in which the subject holds his leg above a chair. Analyses of variance were calculated for all variables, in order to test hypotheses relating to the ability of the tests to

differentiate between neurotic groups. Critical ratios between means of the neurotic groups and that of the normal control group were also calculated where this seemed advisable. Full details will be found in Hildebrand (1953).

II. Results

(i) *Statistical analysis.* Selection of variables for the final factor analysis was based on the following criteria. (1) That all subjects should have completed the test; (2) That the score of the total population on the test should have a normal distribution or a close approximation to it. It will be recalled that it was proposed to exclude the criterion groups from the matrix in order to provide external criteria of I-E and Neuroticism, and the population included in the final matrix was thereby reduced to 95 subjects and consisted of 55 mixed neurotics, 10 obsessionals, 10 reactive depressives and 20 psychopaths. There were in all 22 variables that satisfied the criteria. All scores of the population on the 22 selected variables were transformed into standard scores on a nine point scale. Product moment correlations were then calculated for all variables.

In view of the requirements of the hypotheses and the lack of satisfactory tests of significance which might be applied to the ordinary centroid method of factor analysis (Burt, 1952), it was decided to use Lawley's Maximum Likelihood Method (Thomson, 1950). This technique has the great advantage of applying what Burt calls the best test of significance of extracted factors, namely, a derivation of Bartlett's (1950) method. Lawley's method assumes specific factors; it aims at minimizing the residuals and so accounting for the intercorrelations to the greatest possible extent. 'The method has the property of using the largest amount of information available in the data and gives "efficient" estimates which roughly speaking, are on the average nearer the true values than those obtained by other "inefficient" methods of estimation' (Lawley, 1950). Lawley makes two initial assumptions before using this method. These are, first, that the test scores and the factors of which they are linear functions are approximately normally distributed; secondly, that a hypothesis must be set up as to the number of factors present. Both these assumptions have been met in the present study.

A full centroid analysis was performed on the data to give first estimates of factor loadings and three factors were extracted. The hypothesis was set up that these three factors, which were tentatively identified as 'Neuroticism', 'g' and 'I-E' were sufficient to account for the observed correlations. Two iterations by Lawley's method were carried out as recommended by Lawley and the adjusted factor loadings subtracted from the original matrix.

The resulting residual correlations were tested for significance by Bartlett's (1950) correction of Lawley's test of significance and it was found that there were still significant factors remaining in the matrix. A fourth factor was accordingly extracted. On the basis of this evidence it is suggested that all four extracted factors are significant. Two rotations were sufficient to adjust the factors to psychologically meaningful positions. Full statistical details of these operations will be found by the interested reader in Hildebrand (1953).

(ii) *Interpretation of the factors.* The four factors as finally stabilized accounted for 32% of the total variance. Factor loadings are given in Table 2. The first factor accounted for 12% of the total variance, and was easily identified as a neuroticism factor. As will be seen from Table 2, the four variables with the highest loadings differentiated significantly between normals and neurotics. Moreover, previous studies had shown these tests to be highly saturated in Neuroticism, emotional instability or similar factors.

The second factor extracted was identified as 'g' or general intelligence. This factor accounted for 9% of the total variance. The second factor was confidently identified as an intelligence factor. The first three variables were the three tests of intelligence included in the matrix, and a high negative correlation between intelligence and mistakes and length of time taken on the Porteus maze test seems appropriate.

The third factor extracted was identified as I-E. This factor accounted for 6% of the variance. For the purposes of this table critical ratios were calculated for the scores of the most highly loaded variables between the 45 hysterics and the 45 dysthymics. It will be seen that the tests do in fact discriminate between the groups. For these reasons the factor was identified as an I–E factor.

Table 2. *Final loadings for the four factors*

Variable	No.	Factors I	II	III	IV	h^2
Prog. Matr.	1	179	750	−014	153	618
M.H.V.	2	−036	508	110	119	285
M.M.Q.	3	739*	−006	−257	055	615
Lie-scale	4	−380	050	−371	−009	284
Annoyances	5	429	−273	−012	068	263
S	7	739*	047	−472†	−043	773
T	8	467*	018	099	−142	248
D C	9	831*	−136	057	−127	728
R	10	−246	−236	563‡	−114	446
Power	11	098	744	012	−111	575
Sex humour	12	301	005	222	−198	179
Leg persistence	13	113	−007	−066	167	045
T.D. | Static	15	086	−082	180	520	317
R. } ataxia	16	−041	143	073	513	290
T.D. | Reversals	18	176	−006	052	292	118
R. }	19	021	161	261†	364	227
Body index	20	015	025	−255	035	067
Starting time	21	−167	055	−195	197	107
Total time	22	−062	−307	−071	321	206
Weighted wrong direction	23	257	−423	360†	−114	387
Lifted pencils	24	070	−211	006	098	059
Crossed lines	25	067	−168	425‡	−155	237
Totals		1·220	0·877	0·613	0·502	3·215

* The variables marked thus showed a significant difference $P > 0{\cdot}01$ between normals and neurotics.
† Discriminates between the 45 hysterics and 45 dysthymics $P > 0{\cdot}05$.
‡ Discriminates between the 45 hysterics and 45 dysthymics $P > 0{\cdot}01$.

The fourth factor extracted was identified as arising from the intercorrelations of static ataxia and suggestibility scores. It does not seem possible to identify it more meaningfully.

(iii) *Regression coefficients and factor scores.* The factors having been identified, three deductions remained to be tested. It will be recalled that the experimental design called for discrimination between the various clinical groups by means of measurements made on the basis of scores based on the two factors of Neuroticism and I-E. It was felt that even though the factors were both significant and general they would have no more than a theoretical value unless such a discrimination could be made.

Ledermann's modification of Aitken's Pivotal Condensation method (Thomson, 1950) was used to estimate each individual's score on a given factor. The regression estimates so obtained 'are the best in the sense that they give the highest correlation, taken over a large number of men, between the estimates and the true value of a criterion when the latter can be separately ascertained'. As these requirements had been met in the present

experiment, the scores of all groups were converted to standard scores, the eight highest regression coefficients for the factors Neuroticism and I-E were selected, and the inner products of the standard scores and the regression coefficients calculated to give estimated factor scores for each member of the population on these two factors.

An overall analysis of variance was then made of the estimated scores on the Neuroticism factor, the results of which may be found in Table 3. As will be seen from Table 3, the analysis of variance showed no significant differences among the neurotic population, but a very significant difference when the normal group was included. A *t*-test was carried out between the neurotic population and the normal group and the C.R. equalled 3·74 which is highly significant. This is in correspondence with the experimental results and demonstrates that the battery does discriminate between normals and neurotics at a high degree of confidence.

Table 3

	Anx.	Obs.	Dep.	Hys.	P.P.	Mx.	Total	Normal
No.	25	10	10	25	20	55	145	25
Mean	5·949	5·424	4·647	4·866	5·916	5·364	5·409	3·402
S.D.	1·686	2·068	1·808	1·927	1·803	1·716	1·816	1·686

$F = 1{\cdot}591$ with normals excluded. N.S.
$F = 5{\cdot}858$ with normals included. Significant beyond the 1% level.

Table 4

	Anx.	Obs.	Dep.	Hys.	P.P.	Mx.	Total	Normal
No.	25	10	10	25	20	55	170	25
Mean	1·407	1·082	1·295	2·241	2·637	1·593	1·925	2·875
S.D.	0·986	1·489	1·131	1·271	1·384	1·384	1·399	1·084

$F = 5{\cdot}896$, which is significant beyond the 1% level.

An analysis of variance was also carried out on the factor scores for the I-E factor. Results are found in Table 4. Inspection of the means demonstrates clearly that the distribution follows the prediction. The hysterics and psychopaths are clustered at the extravert end of the distribution, and the dysthymics at the introvert end of the distribution, with the mixed neurotics occupying a central position. The position of the normal group at the extravert end of the I-E continuum does not support 3. 3, but is consistent with the experimental results obtained from this, which suggested that as a whole they could be expected to be highly extraverted. *t*-Tests run between the groups showed no significant differences within the introvert and extravert clusters, but anxiety states were differentiated from hysterics at the 5% level and the other groups at higher levels of significance. Fig. 1 gives a graphic picture of the means of each group in relation to the two axes. The I-E axis has been drawn so as to cut off 10% of the normals and include them in the neurotic group, according to Russell Frazer's (1947) hypothesis that 10% of the normal population is in fact neurotic.

IV. Discussion of Results

Jung's theory of Introversion-Extraversion is essentially descriptive in nature. The procedure used in the present study enabled the writer to support Jung's theoretical formulations on a more experimental basis than had hitherto been possible. It was found possible to divide the experimental population into three categories, which were in accor-

dance with the original hypothesis. These were neurotic introverts, neurotic extraverts, and the normal group.

(1) The introvert category is made up of anxiety states, reactive depressives and obsessionals. They score high on the Neuroticism factor and cluster at the I pole of the I-E factor. Descriptively Eysenck's amalgamation of the three groups into a single category of dysthymics seems substantiated by the experiment. On the factor scores there are no significant differences between the means of the groups making up this category. The factor-score criterion is probably a better one than one made on the basis of one or two tests.

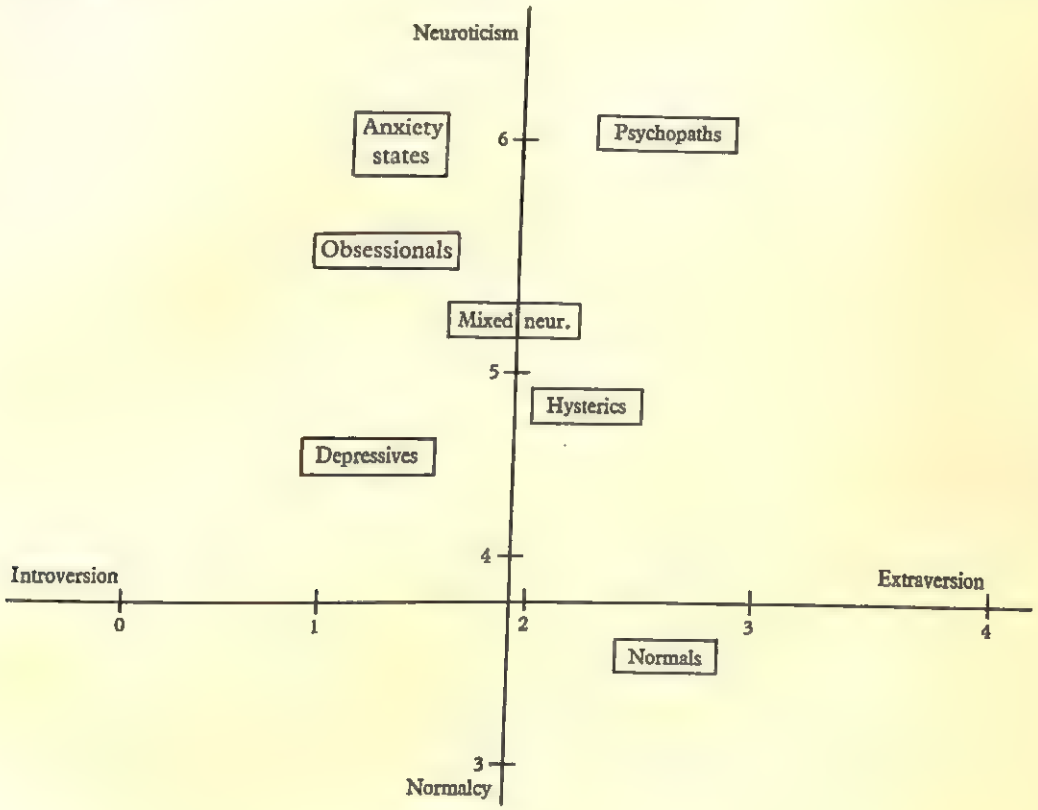


Fig. 1

(2) The extravert category is made up of hysterics and psychopaths. There is no significant difference between the means of the factor scores of these two groups. The actual result of the experiment which finds psychopaths scoring higher on the I-E factor than do the hysterics perhaps needs some discussion. It is inferred that the psychopaths were less inhibited in their test performance than were the hysterics, and it is interesting to note that Mayer-Gross, Moore & Slater (1949) found that on their second canonical variate, which seems to be a measure of I-E, the psychopaths scored higher than did the hysterics.

(3) The normal group did not in fact fulfil the prediction made for it originally. In terms of the other groups it was extremely extraverted although it differed significantly from the neurotic population on the Neuroticism factor. One possible explanation for this extreme degree of extraversion may perhaps be found in Eysenck's 1947 study of

ratings, where he found that such items as 'unskilled', 'narrow interests', and low I.Q. all had significant loadings on what he took to be an I-E factor. These three adjectives describe the control group well, and this may well have contributed to their degree of extraversion. Although it was not possible to test another group on the complete battery, another group of normals was tested on the R scale, which is the highest ranking test on the E pole of the I-E continuum. The control group was found to be significantly more extraverted than this second normal group (Hildebrand, 1956). It is possible that if a less homogeneous and more representative civilian sample of the normal population had been tested on the battery the predictions concerning this group would have been fulfilled. However, there does seem some possibility that the battery may measure I-E within the normal population with perhaps the same efficiency as among the neurotic population.

It is interesting to note that using the rather sophisticated statistical procedures of the present experiment, it was possible to submit Jung's hypothesis to a much more searching test than had hitherto been possible. It is claimed that all the deductions derived from the original theory have been supported at a high degree of significance, with the exception of deduction 3. 3, and reasons have been advanced why this deduction was not confirmed. It should, however, be noted that the entire experiment was purely on a descriptive basis. To quote yet again from Eysenck (1953*c*) 'Factor Analysis is a mathematical procedure which resolves a set of descriptive variables into a small number of categories, components or factors. These factors, themselves, must in the first instance, be regarded as having a purely descriptive function. Under certain circumstances factors may be regarded as hypothetical causal instances underlying and determining the observed relationships between a set of variables. It is only when regarded in this light that they have interest and significance for psychology.'

Whether or not the factors found in this experiment, particularly the I-E factor, may be considered to have hypothetical and causal instances underlying them is not, of course, a question that can be answered on the basis of the data here presented. Yet although the I-E factor as it is found in this experiment seems beyond dispute, it does seem worth while to point out that a certain amount of caution must be used before the factor can be used as a basic postulate in personality research. First, although the experiment was designed to amplify the factor as much as possible, yet the total variance accounted for by the factor amounted to only 6%, only 1% more than that of a factor that it was found impossible to identify satisfactorily. Secondly, of the tests which were suggested as possible measures of I-E, only two differentiated between introverts and extraverts at the 1% level of significance. That, in spite of these results, the battery was as successful as it was is very satisfactory, but the next step in any further investigation of this area of personality would seem to be further descriptive studies rather than a search for the hypothetical causative instance.

This article is based on the writer's unpublished Ph.D. thesis for London University. The writer wishes to thank Dr L. Minski, Physician Superintendent of Belmont Hospital, for kind permission to test patients within the hospital, and Dr M. M. Desai and Mr R. Secker of the Psychology Department, Belmont Hospital, for their many kindnesses. He also wishes to thank his supervising teacher, Dr H. J. Eysenck, for his advice and assistance.

References

Bartlett, M. S. (1950). Tests of significance in factor analysis. *Brit. J. Psychol. Statist. Sect.* **3**, 77–85.

Burt, C. (1952). Tests of significance in factor analysis. *Brit. J. Psychol. Statist. Sect.* **5**, 109–33.

Cattell, R. B. (1946). *The Description and Measurement of Personality.* New World Book Co.

Collier, R. & Emch, M. (1938). Introversion-Extraversion—the concepts and their clinical use. *Amer. J. Psychiat.* **94**, 1045–75.

Eysenck, H. J. (1947). *Dimensions of Personality.* London: Kegan Paul.

Eysenck, H. J. (1953*a*). *The Structure of Human Personality.* London: Methuen.

Eysenck, H. J. (1953*b*). Research Report under contract No. 62558s-56 with the Bureau of Medicine, U.S. Navy. (Unpublished.)

Eysenck, H. J. (1953*c*). The logical basis of factor analysis. *Amer. Psychologist*, **8**, 105–14.

Frazer, R. (1947). *The Incidence of Neurosis among Factory Workers.* London: H.M.S.O.

Freud, S. (1920). *A General Introduction to Psycho-Analysis.* New York: Liverwright.

Freyd, M. (1924). Introverts and Extroverts. *Psychol. Rev.* **5**, 74–87.

Garrett, J. C. M. (1918). General Ability, cleverness and purpose. *Brit. J. Psychol.* **9**, 345–66.

Gross, O. (1902). *Die Cerebrale Sekundarfunktion.* Leipzig.

Guilford, J. P. (1934). Introversion-Extraversion. *Psychol. Bull.* **31**, 331–54.

Heymans, G. & Wiersma, E. (1906). Beitrage zur speziellen Psychologie auf grund einer Massenufersuchung. *Z. Psychol.* **42**, 81–127 *et seq.*

Hildebrand, H. P. (1953). 'A Factorial Study of Introversion-Extraversion by Means of Objective Tests.' Ph.D. Thesis, University of London. (Unpublished.)

Hildebrand, H. P. (1956). On the use of Guilford's introversion-extraversion Questionnaire with a neurotic population. *Rev. Psychol. App.* **6**, 179–185.

Jordan, W F. (1890). *Character as seen in Body and Parentage.* London: Macmillan.

Jung, C. J. (1924). *Psychological Types.* London: Kegan Paul.

Lawley, D. N. (1950). *The Maximum Likelihood Method of Estimation of Factor Loadings.* In Thomson, G.: *The Factorial Analysis of Human Ability.* London University Press. (5th ed.)

Line, W. & Griffin, J. D. M. (1935). The objective determination of factors underlying mental health. *Amer. J. Psychiat.* **91**, 833–42.

Lovell, C. (1945). A study of the factor structure of 13 personality variables. *Educ. Psychol. Meas.* **5**, 335–50.

Mayer-Gross, W., Moore, J. T. N. & Slater, P. (1949). Forecasting the incidence of neurosis in officers in the army and navy. *J. Ment. Sci.* **95**, 80–100.

Minski, L. (1950). *Belmont Hospital, Annual Report.* (Privately printed.)

North, R. D. (1948). An analysis of personality dimensions of introversion-extraversion. *J. Personality*, **17**, 552–67.

Thomson, G. (1950). *The Factorial Analysis of Human Ability.* London University Press. (5th ed.)

Webb, E. (1915). Character and intelligence. *Brit. J. Psychol. Monogr. Suppl.* **1**, No. 3.

(*Manuscript received* 4 *October* 1956)

DETERMINANTS OF EMOTIONALITY IN THE RAT

II. ANTECEDENT FACTORS*

By P. L. BROADHURST

University of London, Institute of Psychiatry (Maudsley Hospital)

A total of 112 rats (56 males and 56 females) is used to investigate the effects of some antecedent factors upon the defecation and ambulation scores derived from Hall's open-field test of emotionality, given under precisely standardized conditions. Conditions of rearing and husbandry are similarly constant, except where systematically varied for experimental purposes. The results show that neither defecation nor ambulation is affected by systematic 'gentling'. Sexual experience does not affect defecation, but increases ambulation significantly, and this effect appears due, among females, to the experience of suckling a litter. Previous experience of the same test at earlier ages decreases defecation among young rats, but not among older ones, while ambulation is unaffected throughout. The effect of ageing itself upon defecation is seen as a decrease among older rats only, young ones being unaffected: conversely, its effect upon ambulation is an increase among young ones only. The previously reported sex difference in response is confirmed; no interaction between this sex difference and any of the experimental variables was detected. It is concluded that these findings further support the status of defecation, as observed in this test, as a measure of an emotional response.

I. Introduction

In a previous paper (Broadhurst, 1957*a*), it was shown that the defecation score in Hall's open-field test (1934) provides a valid measure of emotional response in the rat, and the effect upon it of several variables within the test situation was measured. For this purpose, the previous experience and husbandry of the subjects was carefully kept uniform. In the present paper, the effect of some variables antecedent to the test situation is examined. It is not claimed that this survey of the non-situational determinants of emotionality will be in any way exhaustive; the choice of those reported here has been determined principally by practical considerations relating to further use of the test.

II. Apparatus and procedure

Because it is the independent variables outside the test situation which are to be manipulated in the present investigation, complete uniformity of the open-field test conditions is necessary. Accordingly, the standardization of the test conditions developed in the previous paper (Broadhurst, 1957*a*) was employed throughout, and is briefly specified below. Above a circular arena, 32¾ in. in diameter with white plywood walls 12½ in. high, a battery of loudspeakers and of photographic lamps provided sound ('white' noise) and light fields whose intensity at floor-level averaged 78 db (ref. 0·0002 dynes/sq. cm.) and 165 cp. respectively. (This is the low level of noise used in the first experiment in this series; see Broadhurst 1957*a*, p. 3. This revised determination was made possible by more accurate apparatus which, however, confirms the figure of 94 db previously given for the high level but not now used.) The subjects were transported individually in a covered white enamel pail to the test arena where they remained for 2 min. a day on four successive days. The number of fecal boluses deposited and the number of compartments marked on the floor and roughly equal in size, through which the animal passed, constitute the basis for the defecation and ambulation scores respectively.

Apart from experimental variations in the pre-test experience of the subjects which are the object of investigation and which are described in the appropriate section below, the standard conditions of husbandry used in the previous study (Broadhurst, 1957*a*) obtained. Salient features are the minimum possible handling, together with uniformity of diet and of caging after weaning at 21 days of age until testing took place in the open field as soon after 100 days of age as was practicable.

* This paper is based on part of a thesis approved by the University of London for the degree of Ph.D.

(i) *Introduction*

In the present section the effects of three antecedent variables are considered. The first of these is overt sexual experience. Sexual deprivation, it is thought, may lead to subsequent neurosis (e.g. 'The neurosis owes its existence to a diversion and regression of the libido from the genital position...' (Reich, 1950, p. 131)). While it is not claimed either that sexuality plays as important a part in the experience of the rat as it tends to do in a behaviourally more complex organism like man, or that emotionality as measured by the open-field test can be equated directly with human personality dimensions, it is clear that segregated rats are deprived of a social experience unique in their repertoire of behaviour. Furthermore parturition and/or suckling in females may affect the metabolic balance (Watson, 1905) and thereby alter emotional elimination which is under the control of the sacral branch of the sympathetic nervous system (Hall, 1934).

Practically, it was desirable to know whether or not using our subjects for breeding before the open-field test at the standard 100 days of age would affect their performance thereon. The fact that the social stimulus of another animal in the test situation itself has no effect on scores (Anderson, 1939) suggests that non-situational social experience may be unimportant. The further possibility exists of investigating the effects of pregnancy and parturition among females independently from the effects of suckling a litter, by depriving some subjects of their litters at birth and allowing others to keep them until the usual weaning time, and among the males of investigating any differences between the effect of allowing mating with females alone or in sexual competition with other males.

The second antecedent variable is handling or 'gentling', as it is more expressively termed in American laboratories. It is a common pre-experimental routine in many animal studies, for it renders the subjects more docile, less likely to bite, and reduces flight or startle reactions, including emotional elimination, when touched. The extinction of such behaviour by systematic handling, stroking and petting is therefore of practical convenience in the carrying out of experimentation and is widely employed. However, as Christie (1951) has emphasized, the possible effect of such training on experimental results is frequently ignored. Evidence suggests that it may be important (Bernstein, 1952, 1956; Scott, 1955; Ruegamer, Bernstein & Benjamin, 1954; Levine, Chevalier & Korchin, 1956; Stanley & Monkman, 1956).

With respect to emotionality, Weininger (1954) has shown that gentled rats are significantly more active in the open field on the first day of testing than a group not so treated. There is also a significantly greater weight gain in gentled animals (Weininger, 1954; Bovard, 1954; Weininger, McClelland & Arima, 1954). On the other hand, Hunt & Otis (1955) failed to find differences in defecation scores attributable to gentling, and Gertz's findings (personal communication, 1956) suggest that the effect of gentling may be temporary.

Related to the effect of gentling is that of the third experimental variable to be considered, namely, previous experimental experience. The only systematic investigation appears to be that of Watson (1954*a*, 1954*b*) whose unpublished data show that animals who, after maturity, have undergone training in other experimental situations or have been given the open-field test itself display significantly lowered emotional reactivity when subsequently tested. This demonstration of 'experimentation sophistication', as Watson terms it, is clearly of great practical importance in planning animal studies

involving sequential testing and in interpreting their results. It is desirable, therefore, to see if the effect holds with female subjects and with experience of testing in early life.

(ii) *Experimental design*

The experimental design is shown in Table 1, from which it will be seen that of the three independent variables whose effect on the emotional reactivity scores of rats of both sexes it is desired to establish, that is, sexual experience, handling, and previous test experience, the last two were combined as aspects of one variable, since these two treatments seemed most alike, and hence any possible interactions between them would be of the least interest. This was done because there were insufficient subjects available to allow six replications of a complete fourfold classification in an analysis of variance design.

Table 1. *Experimental design showing treatments imposed before testing*

	Sexual experience		No sexual experience	
Handling	Males*	Females	Males	Females
Previous test experience	Males	Females	Males	Females
Neither above, i.e. a control group with no handling and no previous experience	Males	Females	Males	Females

* The n for each sub-group is 6, total $N = 72$.

(iii) *Subjects and procedure*

Since a decrease in defecation scores was anticipated from the experimental treatments, animals from a high scoring or emotionally reactive strain being developed by selective breeding (Broadhurst, 1957*c*) were used in so far as they were available. The 72 subjects for this experiment thus constituted part of the first generation of the selective breeding study of emotionality. The 60 animals comprising the experimental groups with pre-test treatments were drawn from a number of animals superfluous to that study and control groups without treatment were provided by randomly selecting 12 animals (6 males and 6 females) from among the main populations of animals routinely tested for the purpose of the breeding study. Appropriate t-tests yield non-significant values showing that these control groups impartially represent the main populations in respect of defecation and ambulation scores. As a further check, the possibility of a hereditary bias influencing the scores of the experimental groups was investigated by comparing the scores obtained by the parents of these animals with those of the subjects retained in the selective breeding experiment. In all important respects no significant differences were found. Further details of these checks on the justification for regarding the subjects used as representative of the larger populations from which they came are given elsewhere (Broadhurst, 1956).

Variations from the normal procedures of rearing the animals constitute the experimental treatments and were as follows. (1) *Handling* or gentling followed Weininger's (1954) procedure. Each rat was taken individually from its cage each day and held in the experimenter's lap with his left hand while being stroked from head to tail with his right thumb for 5 min. at a rate of about 90 strokes per minute. This procedure was begun when the 24 animals in this group were 56·5 days old ($\pm 0{\cdot}16$) and continued daily for 57 days until they were given the open-field test. (2) *Previous experience of the test* was ensured by giving the open-field test to the 24 animals in this group, under standard conditions and beginning when they were 53·3 days old ($\pm 0{\cdot}22$). (3) *Opportunity for sexual experience* was allowed simultaneously to all the 36 subjects constituting this group when they were 64 days old ($\pm 0{\cdot}13$). One male was randomly selected from the normal cage complement of three and was mated with 2 females in their home cage. His place was taken by the single female thus displaced from the second cage. Four days later the original caging was restored. This procedure was carried out three times beginning at intervals of 1 week. Thus, the single males (a total group of 6 rats) were always mated with 2 females each, their two cage mates (a total group of 12) were always mated with 1 female per pair. The females concerned, however, were rotated to form these sub-groups. As a result of these matings, all females, except one, became pregnant within 10 days of first mating, and were removed to individual cages before their litters were born. Half of them (9 rats) had their litters augmented to a standard size of 12 pups within 3 days of parturition,

and were permitted to rear them to the normal age of weaning of 21 days, which they all did successfully. The rest of the females constituted the suckling-deprived group, and either were deprived of their pups (7 cases) or lost them (1 case) within 3 days of parturition, or did not become pregnant (1 case).

All animals were tested in the open field under the standard conditions described in section II above when their mean age was 111·4 days (±1·71). There was no significant difference between males and females in this respect.

(iv) *Results*

The mean scores for the various sub-groups, together with a summary of the results of the analyses of variance performed on the two scores, are shown in Table 2. The interactions (three first order and one second order) were uniformly non-significant and are

Table 2. *Mean scores of sub-groups and summary of results of analyses of variance*

		Score				
		Defecation (av. no. of fecal boluses per day)		Ambulation (av. no. of metres per day)		
Source of variance (treatments)	Sub-group	Mean	Significance of F ratio	Mean	Significance of F ratio	D.F.
Sexual experience	Yes	3·2	N.S.	8·7	$P<0·01$	1,60
	No	2·6		6·9		
Other treatments	Handling	3·3	$P<0·01$	8·8	N.S.	2,60
	Previous testing	2·1		7·0		
	Neither	3·3		7·6		
Sex	Males	2·8	N.S.	7·1	$P<0·05$	1,60
	Females	3·0		8·5		

Table 3. *Significance of differences in defecation scores between different treatments*

Comparison	Mean difference	Significance
Previous testing *v.* Handling	1·2	$P<0·01$
Previous testing *v.* Neither treatment	1·1	$P<0·01$
Handling *v.* Neither treatment	0·1	N.S.

therefore omitted from the table. The significance of the F-ratio associated with the main effect upon defecation scores designated 'other treatments' permits the application of t-tests to evaluate the significance of the differences (one-tail tests of significance were used where appropriate) between the defecation scores of the three sub-groups constituting this variable. The results are given in Table 3. In a further subsidiary analysis of the group allowed sexual experience t-tests were used to investigate the differences between the two groups given slightly different sexual experience, since these sub-groups could not have been conveniently differentiated in the main analysis (Table 2). The outcome is shown in Table 4.

These results clearly fail to confirm the reported decrease in emotionality following gentling, despite its effect upon growth. The effect upon weight—both males and females at testing were heavier than ungentled animals—was not, however, significant. The only experimental variable which has had an effect upon the defecation scores is that of previous experience of the test. Animals which have been given the test once before at the age of some 53 days defecate less than those undergoing it for the first time. (The effect of re-testing later in life will be discussed in the next section.) The usual sex difference in defecation scores is not seen, probably because the emotionally reactive strain, from

Table 4. *Significance of differences between groups given slightly different sexual experience*

Sex	Treatment	n	Score: Defecation (av. no. of fecal boluses per day) Mean	Significance of difference	Score: Ambulation (av. no. of metres per day) Mean	Significance of difference
Males	Mated in pairs with single female	12	3·0	N.S.	8·4	N.S.
	Mated alone with two females	6	2·6		7·7	
Females	Allowed to suckle litter	9	3·4	N.S.	10·4	$P<0·01$
	Deprived of litter	9	3·6		8·1	

which most—54 out of 72—of the subjects were drawn, does not show this difference as clearly as does the non-emotional strain, or the unselected populations used in the previous study (see Broadhurst, 1957*c* and 1957*a*, respectively). This interaction between strain and sex differences in selectively bred animals is fully discussed elsewhere (Broadhurst, 1958): there seems no reason to suppose it has any bearing on the outcome of the present experiments.

The ambulation scores, which do not in general display this interaction between sex and strain, show the usual sex differences, the females being significantly more active than the males. In contrast with its negligible effect upon defecation scores, the sexual experience variable has an important effect upon ambulation. With respect to the females' experience, it is possible to be precise about which of the inter-group differences is associated with this effect, since the females without sexual experience have a mean ambulation score of 7·8 m. per day, which is insignificantly different from that of the litter-deprived group of females (Table 4). This suggests that it is the experience of suckling a litter which is crucial.

III. Effects of age

(i) *Introduction*

This section is concerned primarily with the effect of age upon open-field test scores. Small variations in age of testing mature animals have no effect (Broadhurst, 1957*a*), but O'Kelly's results (1940) suggest that young rats show a decrease in defecation scores with age of testing during early life. The most extensive analysis of the effects of age on scores of mature (100 days and older) rats is provided by Watson's unpublished data (1954*b*). For example, his results from four groups of males showed a systematic increase of defecation scores with advancing age from 131 to 341 days, and the difference between the extreme groups reaches the 3% level of significance. It is proposed to attempt to extend these findings, first, by including both immature animals and mature ones of an age intermediate between Watson's extreme groups, and secondly, by using both sexes in the experimental analysis. In view of the highly significant findings relating to previous test experience among young rats and reported in the previous section, the opportunity will be taken to examine the effect of such experience upon testing much later in life.

In effect, these objectives were accomplished by dividing the investigation into two parts, and investigating first the effects of ageing before the standard age of testing had

been reached, and secondly, the effects of ageing after maturity. Since the previous work would lead one to predict a decrease of emotional defecation with age in young rats and the reverse in mature ones, animals from the high-scoring, or emotionally reactive strain, derived from the selective breeding experiment mentioned earlier, were used for the first part, and animals from the low-scoring, or non-reactive strain for the second. In this way the expected effects could be most clearly shown.

(ii) *Subjects and procedure*

All subjects were reared and tested under the standard conditions described earlier (section II). The only variations—which constitute the independent variables to be studied in this section—lay in the testing, or in some cases re-testing, at ages systematically different from the standard one of around 100 days.

For the first part of the experiment, the subjects were the 12 male and 12 female rats of the emotionally reactive strain who constituted the 'previous test-experience group' in the previous section. For that purpose they were tested at a mean age of 53·3 days before being re-tested 60 days later at a mean age of 113·3 days (±0·22).

The subjects for the second part of this experiment were 10 male and 10 female rats of the second generation of the emotionally non-reactive strain whose open-field testing was delayed until a mean age of 242·3 days (±0·17). For comparison, control groups also of 10 male and 10 female non-emotional rats were selected. Their mean defecation and ambulation scores did not differ significantly (by *t*-test) from those of the main population, also tested at the standard age. These control groups were given the test for the second time at a mean age of 237·6 days (±0·60).

(iii) *Results*

The outcome of the comparisons relating to the younger animals is presented in Table 5. Because the handling variable is the only one in which the group tested at 50 days of age differs from all the rest of the subjects used in the experiment reported in the previous section—numbers of each sex, and of those allowed sexual experience being proportion-

Table 5. *The effects of ageing on the test scores of young rats*

Score	Approximate age at 1st testing		Significance of difference,	Approximate age at 2nd testing	Significance of difference,	Significance of difference, 1st *v.* 2nd testing at 100 days
	50 days	100 days	50 *v.* 100 days	(100 days)	50 *v.* 100 days	
Defecation (av. no. of fecal boluses per day)	3·1	3·3	N.S.	2·1	$P<0·01$	$P<0·01$
Ambulation (av. no. of metres per day)	6·0	8·3	$P<0·001$	7·0	$P<0·001$	N.S.

ately equal—and because handling has been shown to have had no effect, these latter subjects therefore form an appropriate control group for the results of testing at the standard age (see col. 3). The differences for each score, evaluated by *t*-test, are shown in the fourth column. The rest of the table shows the effect of re-testing, that is, a comparison is made of the scores returned at 50 and 100 days of age by the same subjects, these differences being accordingly evaluated by Sandler's (1955) *A*-test for correlated measures (see col. 6). They should be compared with those between naïve and test-sophisticated rats of the *same* age derived from the analyses of the previous section (Tables 2 and 3) and which are reproduced for convenience in the final column.

The results of the comparisons relating to the older age groups are presented in similar form in Table 6. The significance of the difference between the scores of the same animals

on the two occasions of testing was evaluated as before with the *A*-test (col. 6); otherwise double classification analyses of variance were used in which the sex difference was included as the second variable. The significance of the *F*-ratios associated with the main effects due to age of testing or the previous test experience—the other variable in the analysis of variance—are indicated in the table (cols. 4 and 7 respectively), but, as might have been expected from these subjects belonging to the emotionally non-reactive strain (Broadhurst, 1958), the sex differences were significant throughout and in the usual direction—males defecating more than females and ambulating less—and are not therefore shown in the table. There were no significant interactions between age and sex.

Table 6. *The effects of ageing upon the test scores of older rats*

Score	Approximate age at 1st testing		Significance of difference, 100 *v.* 240 days	Approximate age at 2nd testing (240 days)	Significance of difference, 100 *v.* 240 days	Significance of difference, 1st *v.* 2nd testing at 240 days
	100 days	240 days				
Defecation (av. no. of fecal boluses per day)	2·6	1·5	$P<0{\cdot}05$	1·3	$P<0{\cdot}001$	N.S.
Ambulation (av. no. of metres per day)	7·8	7·6	N.S.	7·4	N.S.	N.S.

Trends in open-field test scores with age and test experience can be discerned in these results, and the comparison of the first test and the re-test at the different ages permits a rough analysis of the influence attributable to these two factors, which may be stated as follows. First, it is clear that there is little change in defecation scores between 50 and 100 days, whereas between 100 and 240 a significant decline occurs. But, while testing at 50 days causes a significant decrease on re-testing at 100 days, testing at the latter age has no additional influence on the decline in defecation scores at 240 days. Second, though ambulation scores increase significantly between 50 and 100 days of age, they change little between 100 and 240 days, and the effect of prior testing at either 50 or 100 days upon later performance is negligible. There is no evidence suggestive of sex differences in any of these respects, but it must be stressed that these findings may have limited generality because animals with characteristically high or low defecation scores were used in the first and second parts of this study respectively, and it has not been established that different strains of animals do not vary differently with age. Indeed the fact that the present findings run counter to those of both O'Kelly (1940), and Watson (1954*b*) suggests that such may be the case, though differences in testing technique may also be important.

V. Discussion and conclusions

The findings from the experiments reported in this paper suggest the following conclusions: the only antecedent variables shown to be indubitably related to emotional reactivity as measured by defecation scores are prior experience of the same test, and age after maturity. With respect to others investigated, it is, of course, only possible to say that their effect, if any, has not been detected under what seem to be favourable conditions. Thus, further work, reported elsewhere (Broadhurst, 1956), shows, first, that depriving rats of form vision by screening the front of their cages with opaque Perspex throughout their life before testing has no effect on open-field test scores, and secondly,

that such scores are not related to rodent body-build, as measured by Watson's (1957) index.

The confirmation of the test-experience effect described by Watson (1954*a*), though partial—it was found to be present among young rats but absent in ones of a similar age to his subjects—strengthens the conclusion that previous test experience does lower defecation scores on subsequent testing. The explanation is not far to seek: adaptation to the mild stress situation represented by the open-field test is clearly demonstrated by the well-established decrease in defecation scores with repeated testing on successive days—indeed it is this extinction which supports the validity of defecation as a fear response. Clearly it may be anticipated that part of such adaptive learning derived from one series of daily tests will be carried over to a subsequent series, thus giving a reduced score. The fact that learning situations different from the open-field test have a similar effect (Watson, 1954*b*) may be expected from the generality of the defecation response to strange situations (Parker, 1939).

But it is this generality which would lead one to expect that animals, gentled so that they—*inter alia*—no longer defecate when handled, would not do so when faced with the open-field situation, whereas our findings fail to confirm those of others which suggest the effect mentioned. However, gentling in most of these cases was started at an earlier age than that used here, and it may well be that 'critical periods' of susceptibility to certain kinds of learning occur in the life of the rat, just as they have been shown to do in other species (Scott & Marston, 1950; Tinbergen, 1951, p. 150). It is clear that some such hypothesis relating to the effects of handling upon the increase in body-weight in rats would reconcile the finding that 21 days of gentling begun at age 23 days increases adult body-weight significantly (Weininger, 1954), whereas 57 days of gentling begun at age 56 days did not. Similarly, the result that, among our subjects, prior test experience has a potent effect if given at 50 days, but not if given at 100 days of age accords with this view.

It has thus been shown that the defecation scores derived from the standardized open-field test used here are insensitive to variations in several kinds of antecedent factors which might have been expected to affect them. These findings consequently further enhance the usefulness of the measure as an index of an emotional response. Only previous experience of the test and large differences in age after maturity have been shown to be important in this respect, and the effect of these variables, like that of the sex difference previously reported (Broadhurst, 1957*a*, *b*) and, in general, confirmed here, must be taken into account in making practical use of the test.

The ambulation scores represent a less important index of emotionality, since they largely reflect the intensity of evocation of exploratory tendencies (Broadhurst, 1957*a*). Consequently they are of less interest in the present connexion. The findings differ, however, from those relating to defecation in two principal ways. First, there is no evidence that ambulation is influenced at all by previous testing. This suggests that the exploratory drive is easily aroused but is weak, since learning based upon its reduction is impermanent. Secondly, there is a marked effect attributable to previous experience of a sexual nature. This can be interpreted in several ways. The absence of corresponding changes in defecation scores argues against its being due to a decrease in fearfulness, while the specificity of the causative factor attributable in females—the experience of suckling a litter—suggests a change in activity level, perhaps learned, or perhaps metabolically

determined. Why the experience of mating should similarly affect males is not apparent from these data: what is clear is that further work would be needed to confirm and elucidate these particular findings.

References

Anderson, E. E. (1939). The effect of the presence of a second animal upon emotional behavior in the male albino rat. *J. Soc. Psychol.* **10**, 265–8.

Bernstein, L. (1952). A note on Christie's: 'Experimental naïveté and experiential naïveté'. *Psychol. Bull.* **49**, 38–40.

Bernstein, L. (1956). A reply to Scott. *J. Abnorm. Soc. Psychol.* **53**, 141–2.

Bovard, E. W. (1954). A theory to account for the effects of early handling on viability of the albino rat. *Science*, **120**, 187.

Broadhurst, P. L. (1956). *Emotionality in the Rat: its Determinants, Inheritance, and Relation to some Aspects of Motivation.* Ph.D. thesis, University of London.

Broadhurst, P. L. (1957*a*). Determinants of emotionality in the rat: I. Situational factors. *Brit. J. Psychol.* **48**, 1–12.

Broadhurst, P. L. (1957*b*). Determinants of emotionality in the rat: III. Strain differences. *J. Comp. Physiol. Psychol.* (In the Press.)

Broadhurst, P. L. (1958). Emotionality in the domestic rat as modified by selective breeding. (In the Press.)

Christie, R. (1951). Experimental naïveté and experiential naïveté. *Psychol. Bull.* **48**, 327–39.

Hall, C. S. (1934). Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. Psychol.* **18**, 385–403.

Hunt, H. F. & Otis, L. S. (1955). Restricted experience and 'timidity' in the rat. *Amer. Psychologist*, **10**, 432. (Abstract.)

Levine, S., Chevalier, J. A. & Korchin, S. J. (1956). The effect of early shock and handling on later avoidance learning. *J. Personality*, **24**, 475–93.

O'Kelly, L. I. (1940). The validity of defecation as a measure of emotionality in the rat. *J. Gen. Psychol.* **23**, 75–87.

Parker, M. M. (1939). The interrelationship of six different situations in the measurement of emotionality in the adult albino rat. *Psychol. Bull.* **36**, 564–5. (Abstract.)

Reich, W. (1950). *Character Analysis* (3rd ed.). London: Vision Press.

Ruegamer, W. R., Bernstein, L. & Benjamin, J. D. (1954). Growth, food utilization, and thyroid activity in the albino rat as a function of extra handling. *Science*, **120**, 184–5.

Sandler, J. (1955). A test of the significance of the difference between the means of correlated measures, based on a simplification of Student's *t*. *Brit. J. Psychol.* **46**, 225–6.

Scott, J. H. (1955). Some effects at maturity of gentling, ignoring, or shocking rats during infancy. *J. Abnorm. Soc. Psychol.* **51**, 412–14.

Scott, J. P. & Marston, M. (1950). Critical periods affecting the development of normal and maladjustive social behavior of puppies. *J. Genet. Psychol.* **77**, 25–60.

Stanley, W. C. & Monkman, J. A. (1956). A test for specific and general behavioral effect of infantile stimulation with shock in the mouse. *J. Abnorm. Soc. Psychol.* **53**, 19–22.

Tinbergen, N. (1951). *The Study of Instinct.* Oxford: Clarendon Press.

Watson, J. B. (1905). The effect of the bearing of young upon the body-weight and weight of the central nervous system of the female white rat. *J. Comp. Neurol. Psychol.* **15**, 514–24.

Watson, R. H. J. (1954*a*). The effects of age and experience on the performance of the white rat on a test of emotionality. *Bull. Brit. Psychol. Soc.* **23** (inset), 6. (Abstract.)

Watson, R. H. J. (1954*b*). *The Behaviour of the Rat in Conflict Situations Presented Serially.* Ph.D. thesis, University of London.

Watson, R. H. J. (1957). A factor analysis of body measurements in the rat. (In preparation.)

Weininger, O. (1954). Physiological damage under emotional stress as a function of early experience. *Science*, **119**, 285–6.

Weininger, O., McClelland, W. J. & Arima, R. K. (1954). Gentling and weight gain in the albino rat. *Canad. J. Psychol.* **8**, 147–51.

(*Manuscript received* 12 *November* 1956)

SURPRISE, ISOLATION, AND STRUCTURAL CHANGE AS FACTORS AFFECTING RECALL OF A TEMPORAL SERIES

By ROBERT T. GREEN

Psychology Department, University College, London

Summary

Two experiments are considered which throw light on certain aspects of what has become known as the von Restorff effect. The first is designed to find out whether items in a serial learning task are favoured in recall because unexpected change attracts attention. The results show that change as such, expected or unexpected, has an effect, though unexpected change does have a somewhat greater effect on recall scores. Subjective reports suggest that this is due, at least in part, to the fact that when the instructions are understood they serve to direct attention. The second experiment directly attacks the question of whether 'isolation' has any distinct effect on probability of recall, and demonstrates that structural change in a temporal series is the crucial variable.

Experiment 1. The effect of instructions intended to reduce the 'surprise' value of structural change

I. *Introduction*

In a previous paper (Green, 1956) evidence was considered which indicated that an unexpected change in structure of a temporal series increased the probability of recalling the item introducing the change. The degree of 'surprise' occasioned by a particular change was inferred by assuming that subjects' expectations were formed by an inductive process. Thus a particular nonsense syllable occurring in the fourth serial position was more likely to be recalled if it were preceded by three items of another sort (in this case three-digit numbers) than if it were preceded by three other nonsense syllables. Later changes of this kind in the same list would be less unexpected since once the uniform series is broken the inductive principle would lead to a recognition of the possibility of similar changes. Put more simply: after the first change the subject would know he was dealing with a mixed list. The data seemed to support this contention inasmuch as the first change showed a large differential recall effect whereas later changes produced progressively smaller differences in recall scores.

In order to verify this 'surprise' hypothesis more directly it was decided to run two more groups of subjects under similar conditions to those holding for half the subjects in the original experiment, with the exception that the 46 new subjects were given detailed information about the structure of the list immediately prior to presentation. It was predicted that, if surprise is an important factor and if the information reduces or eliminates surprise, then the item introducing the change would be recalled less often by the subjects with knowledge of structure than those subjects without such knowledge. In fact it should be recalled no better than by the groups for whom the item introduced no change.

II. *Experimental procedure*

The experiment consisted in displaying 20 items consecutively for 5 sec. each and requesting the subjects to recall in writing as many of the items as possible without regard to order. The procedure was identical with that described in the original paper except that in the instructions given to 23 of the subjects the

following addition was made. 'Finally, let me say something about the way the list is made up so you will have some idea of what to expect. In all there are 20 items, one on each card. I shall expose these cards one at a time for a period of 5 sec. each. On half of the cards there is a nonsense syllable. The other ten carry three-digit numbers. A nonsense syllable is a vowel with a consonant either side such as *BOK*. This particular nonsense syllable will not appear in the list, but 10 of the items will be of a similar nature. An example of a three-digit number would be *895* which again will not appear, although the other 10 items will be of this sort. The items are presented in a particular order. The first three will be nonsense syllables, then a number, followed by six more nonsense syllables, then six numbers, another nonsense syllable, and a final three numbers.' A similar addition was made for the other 23 subjects except that 'nonsense syllable(s)' became 'number(s)' and vice-versa.

So the 46 subjects given the lists *a* or *A* in the original experiment were matched by two groups of the same sizes on the same lists. For the sake of convenience, the lists *a* and *A* are shown again in Table 1. As may be seen, the crucial item in position *4*, which for group *A* is *406* and for group *a* is *HOF*, is surrounded by items of the other sort. The sample consisted mainly of university graduates and students with the addition of a few experimentally naïve adults from other sources to round off the number in each group.

Table 1. *Structure of the two lists*

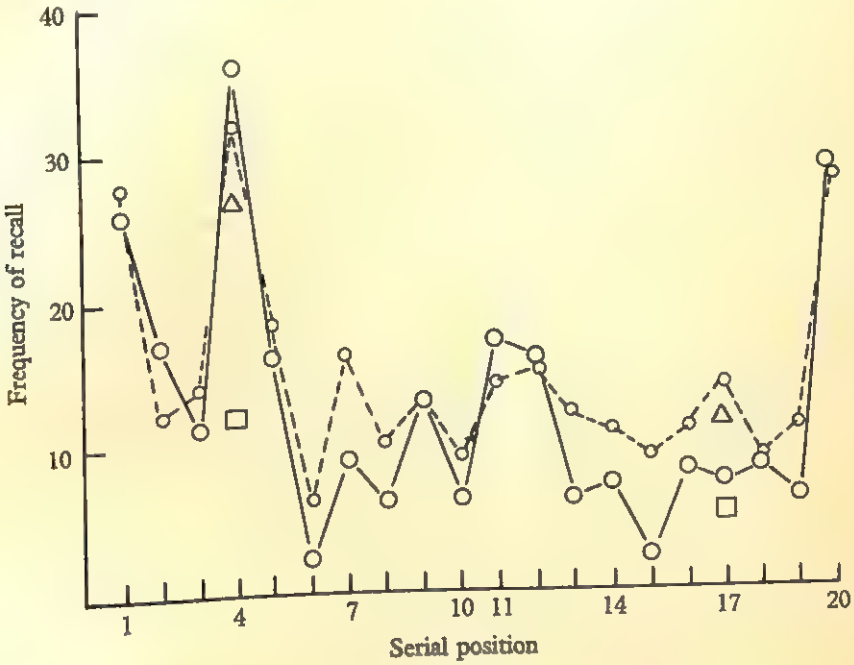
Serial position	List *A*	List *a*
1	*GUB*	20
2	*KEV*	19
3	*DAC*	18
4	*406*	17
5	*TER*	16
6	*WAJ*	15
7	*SIH*	14
8	*RUL*	13
9	*VOM*	12
10	*FIP*	11
11	*417*	10
12	*269*	9
13	*341*	8
14	*562*	7
15	*738*	6
16	*179*	5
17	*HOF*	*4*
18	*258*	3
19	*763*	2
20	*581*	1

Interest in the results was mainly centred on the probability of recall of item *4*. It was intended simply to compare these recall probabilities for the two conditions, with and without knowledge of structure, but due to an unforeseen difficulty a more complex analysis proved necessary.

III. *Results*

The raw data are summarized in Fig. 1. Although item *4* is recalled less often for the *K* (with knowledge) group, the difference—32 subjects against 36 who recalled this item under *I* (ignorance) conditions—is so small that its significance is lost. With these data the temptation is to accept the null hypothesis and assume that knowledge of structure has no effect. However, an odd point brought to light an important factor, which on reflexion seems obvious but which had nevertheless been overlooked. Turning our attention to the isolated item in the list in position *17* we see that the *K* group recalled this not less often but twice as often as the *I* group. In position *17* we should anticipate practically no difference, since in the original experiment the effect of isolating an item in this position was minimal (recalled by 7 subjects when isolated, and by 5 when massed).

In seeking for an explanation of this anomaly we are obliged to query the tacit assumption that the two groups were in fact comparable. Sampling errors are always with us, and in this case if we look at the total recall figures we see that while the *I* group recalled altogether 248 items, the *K* group recalled a total of 292. Testing this difference in a 2×2 table which includes the forgetting frequencies we find that it is significant at the $2\frac{1}{2}\%$ level on a two-tail test.



																					Total
I group	26	17	11	36	16	2	9	6	13	6	17	16	6	7	2	8	7	8	6	29	248
K group	28	12	14	32	18	6	16	10	13	9	14	15	12	11	9	11	14	9	11	28	292
K corrected	23·8	10·2	11·9	27·2	15·3	5·1	13·6	8·5	11·0	7·6	11·9	12·7	10·2	9·3	7·6	9·3	11·9	7·6	9·3	23·8	248

Fig. 1. Effect of knowledge of structure. O—O, ignorant of structure; o - - - o, with knowledge of structure; □, massed; △, corrected.

If there is a tendency for overall better recall by the *K* group then we should expect this to reduce the difference for item *4*. The easiest way to allow for this factor of overall recall probability is simply to reduce all scores for the *K* group in the ratio 248:292. The result of this is to reduce the score for the *K* group in position *4* from 32 to 27·2. Incidentally, the recall score for the other group in the original experiment for whom the item in position *4* was massed, is raised slightly from 12 to 13·5 because their overall recall score was low; 221 in fact. Corrected values appear in Fig. 1 as triangles. If we compare the recall scores 36 and 27·2 for the *K* and *I* groups we find that the difference is significant at the 5% level on a two-tail test. Two objections need to be met before we can reject the null hypothesis with this level of confidence. First it may be argued that the groups differ in overall recall scores not because the samples were unfortunate, but *because* of the difference in treatment accorded them. This is not altogether fantastic, since knowledge of structure could decrease anxiety and lead to a more composed approach to the task. Even if this were true it would still be irrelevant for the purpose of testing the 'surprise' hypothesis. Suppose, to use a *reductio ad absurdum* argument, that

overall scores for the K group were increased to the point where, although item 4 no longer stood out in the serial position curve, it was nevertheless recalled in absolute terms more often than for the I group, we would still have proved the point that the better recall due to unexpected change is lost under K conditions. It is the relative and not the absolute recall score of item 4 that is of interest in assessing the surprise effect. Relative, that is, to the remaining items.

Furthermore, when we consider that the I sample came from the students of an evening institute, whereas the K sample was of university students and graduates, it is hardly surprising to find a higher recall ability in the second sample. Allowing then, that some sort of adjustment must be made for overall recall scores in order to make a comparison of the type we are interested in, are we justified in making a simple proportional correction throughout? We are, provided we are willing to make the assumption that all items are affected in the same proportions by any such overall effect. Although this seems a reasonable enough assumption to make, it may still be wrong. To guard against this danger another method of analysis is required.

Table 2. *Frequencies of recalling item 4 as a function of total number of items recalled*

	I Group			*K* Group				
P	Including 4	Forgetting 4	Total items recalled	Including 4	Forgetting 4	*P*	r_b
1·00	1	0	1	0	0	—	—
0·50	1	1	2	0	2	0·00	+0·577
1·00	2	0	3	2	2	0·50	+0·500
0·77	10	3	4	2	2	0·50	+0·251
0·50	5	5	5	2	0	1·00	−0·378
1·00	6	0	6	7	5	0·58	+0·439
1·00	4	0	7	6	2	0·57	+0·316
1·00	3	0	8	7	1	0·88	+0·194
0·67	2	1	9	3	0	1·00	−0·448
1·00	2	0	10	2	0	1·00	0·000
—	0	0	11	1	0	1·00	—
Total subjects	36	10		32	14	Σr_b	+1·451
	(*a*)	(*b*)		(*c*)	(*d*)	$\bar{r}_b$	+0·181
						Σr_b^2	1·320

$$r_b = \frac{(ad - bc)}{\sqrt{(a+c)\,(b+d)\,(a+b)\,(c+d)}}$$

$$t = \frac{0{\cdot}181\sqrt{8}}{\sqrt{0{\cdot}165 - 0{\cdot}033}} = 1{\cdot}415 \qquad \eta = 8 \qquad P \simeq 0{\cdot}1 \text{ (one-tail)}$$

Although the groups are not equated in terms of overall recall scores, and it is impossible in principle to know beforehand whether or not they will be so equated, we can still overcome this difficulty by equating them along this dimension using the actual data. For example, we can take all the subjects of the K group who recalled a total of five items correctly and compare them with those of the I group who scored the same total. Our hypothesis is that a greater proportion of the I group than the K group will have recalled item 4.

An analysis of the data for all actual scores is given in Table 2. Comparisons for totals of 1 and 11 are impossible because no member of the K group recalled only one item and no member of the I group recalled as many as 11 items, so we lose two subjects, one from either group, by restricting our analysis to the common range.

A graph plotting the recall probabilities of item 4 against total items recalled for each total from 2 to 10 (Fig. 2) is too erratic to tell us much because of the small numbers

occurring in some of the cells. A moving average, grouping in fours, gives a smoother picture (Fig. 3). As predicted, the curve for the *I* group lies above that for the *K* group.

In order to make a valid statistical test the following procedure was adopted.* For each total in the common range the statistic τ_b was calculated. The formula is given in Table 2 and a full discussion is to be found in Kendall (1948). Without going into detail

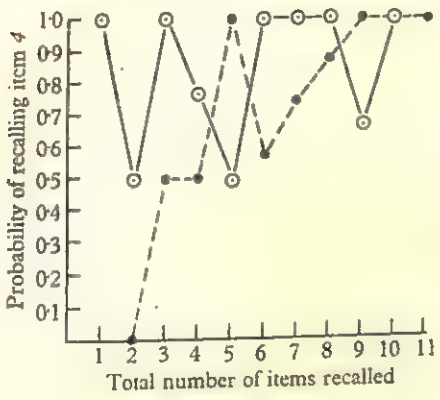


Fig. 2

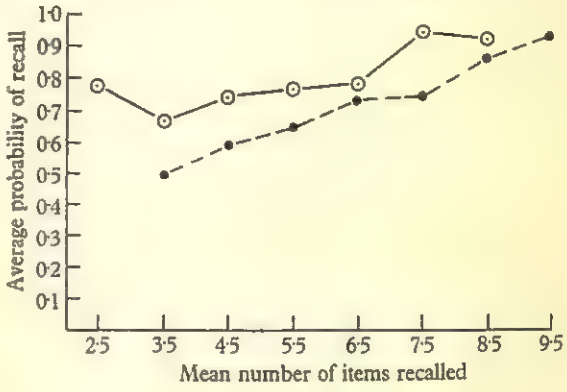


Fig. 3

Fig. 2. Recall probabilities of item *4* as a function of total number of items recalled. ⊙—⊙, *I* group: •- - -•, *K* group.

Fig. 3. Moving average over four sub-groups. ⊙—⊙, *I* group; •- - -•, *K* group.

we may note that τ_b will be positive when the graph for the *I* group in Fig. 2 lies above that for the *K* group, and negative when the opposite trend occurs. This measure is also partly weighted to allow for the number of subjects contributing to each sub-sample. So in essence we have a means of testing whether the curve for the *I* group lies above that for the *K* group. Testing against a theoretical mean of zero, a *t*-test gives $P \simeq 0{\cdot}1$ (one-tail), which is not very impressive. From a technical point of view a more satisfactory treatment is to use the Fisher-Yates exact test for 2×2 tables for each total compared, and combine the probabilities. This method gives $P \simeq 0{\cdot}07$.

IV. *Discussion*

Although the predicted reduction in the recall probability of the item in position *4* was found when the effect of overall recall scores was allowed for, the drop was by no means as marked as anticipated. If 'surprise' is the important factor and if previous instructions eliminate it then it is to be supposed that the recall of the item in position *4* under the *K* condition would be very similar to that where the same item is not isolated under the *I* condition. The actual difference between 32 and 12 is highly significant, and even if the corrected scores of 27·2 and 13·5 are used $\chi^2 = 8{\cdot}27$ and $P \simeq 0{\cdot}004$.

Several possible explanations seem to offer themselves apart from sampling errors. The anomaly of the result was apparent some time before all the subjects had been run, and advantage was taken of this fact by asking the remaining subjects for retrospective reports. Two important points arose from these inquiries. In the first place the instructions were too complex to be grasped in detail and although most subjects realized that

* For the original idea behind this analysis I am indebted to Prof. L. J. Cronbach, and for the actual statistical technique to Dr A. R. Jonckheere and Dr R. Audley.

Bureau Edni. Psy. Research

two types of material would appear the actual order was not mastered. Secondly, those subjects who grasped the fact that an item would be isolated at either end of the list reported waiting for these items to appear. In other words, the more effective the instructions were in eliminating 'surprise' the more they served to direct attention to the isolated items.

So, it would seem, *change* as such has attention-getting value whether expected or not, though the difference between the *K* and *I* groups suggests that unexpected change has additional attention-getting value. In the realm of sensory perception it is almost trite to refer to the effectiveness of change in securing attention, even to the extent of not noticing the tick of a clock until it stops. It is hardly surprising that in an experiment of this type the greater attention accorded an item introducing a change is reflected in a greater probability of its recall.

Experiment 2. The relative effects of isolation and structural change upon recall

I. *Introduction*

It has been suggested by von Restorff (1933) that, in a temporal series, those items which are isolated, in the sense of being preceded and succeeded by items of a different type, will be better remembered because the figure-ground effect holds even when the material is presented item by item. It is supposed that this is in some way due to changes in the engrams, similar engrams becoming confused while the odd engram escapes.

More recent work by Green (1956) has shown that structural change is the more influential factor in affecting recall probabilities. In a list containing two critical items, only the first one is remembered significantly better when isolated than when surrounded by similar material. The suggested explanation was that while the first isolated item introduced a structural change, being the first item of a new type, the mixed nature of the list was well established by the time the second isolated item appeared. However, there was still a weak tendency for the second critical item to be remembered better when isolated than when massed, and this leaves us with the question as to whether this is a residual effect of structural change, or whether it indicates a small but genuine isolation factor. To settle this question it is necessary to devise an experiment which will separate these two factors.

II. *Experimental procedure*

In order to distinguish between the effects of isolation and structural change three groups were used, each being presented with a list that differed in structure, although the same 12 items were involved. These three lists are shown in Table 3. To control serial position and item-uniqueness effects the analysis is confined to the probability of recalling *CZ*, the item which occurs in the fourth serial position for all three groups. A perceptual paradigm of noughts and crosses is put alongside the lists to make the structural properties clearer.

In the change group item *4* is the first letter combination to appear and thus introduces a structural change. For the isolation group, item *4* not only introduces a structural change but is in addition isolated. So if isolation, as distinct from mere structural change, does facilitate recall then *CZ* should be remembered more often by the isolation group than the change group. This boils down in effect to asking whether or not the *succeeding* items have any influence upon the probability of recalling the critical item. If no difference is found between the two experimental groups, then in behaviourist serial interference terms, such as developed by Gibson (1940), the hypothesis becomes that while pro-active interference is considerable, retro-active interference is negligible.

The control group was added to provide a datum for the two experimental groups. In this list *CZ* is neither isolated nor is it the first letter item. It is predicted that both experimental groups will recall *CZ* better than the control group. If the structure of the succeeding items is irrelevant then there should be no difference between the two experimental groups.

The 167 subjects were students at a Technical College, nearly all males about 16 years of age. The material was presented to groups of 15–20 at a time by means of film strip and a projector in a semi-darkened room. Each item appeared for 3 sec. and was followed immediately by the next. Before the experiment proper started an informal chat put the subjects at their ease. They were assured that this was not an intelligence nor a personality test, and complete anonymity was guaranteed. Then formal instructions were given as follows: 'This is a memory experiment. I shall show you a series of items one at a time. After the series is complete you will have 1 minute in which to write on the slip provided as many of the items as you can recall. The order makes no difference. I am interested only in the total number of items that you recall correctly. Please attend carefully and avoid making any comments. Is there anything you are not quite sure about?' Nobody appeared to require more than the minute in which to record all they could remember.

Table 3. *Structure of lists*

Serial positions	Control group		Change group		Isolation group	
1	*17*	×	*17*	×	*17*	×
2	*RQ*	0	*53*	×	*53*	×
3	*94*	×	*94*	×	*94*	×
4	*CZ*	0	*CZ*	0	*CZ*	0
5	*68*	×	*RQ*	0	*68*	×
6	*TJ*	0	*TJ*	0	*21*	×
7	*79*	×	*BD*	0	*79*	×
8	*NL*	0	*NL*	0	*RQ*	0
9	*53*	×	*XP*	0	*TJ*	0
10	*BD*	0	*68*	×	*BD*	0
11	*21*	×	*21*	×	*NL*	0
12	*XP*	0	*79*	×	*XP*	0

III. *Results*

To give a general picture serial position curves for the three groups are shown in Fig. 4. A fair approximation to the usual bow shape is apparent except for position *4* in the case of the two experimental groups. While only 30% of the control group recalled *CZ*, some 60% of the isolation group and 65% of the change group recalled this critical item. The actual recall frequencies are given in Table 4 and apparently the difference between the control and experimental groups is highly significant. The difference between the two experimental groups is not marked and occurs in the opposite direction to that predicted by the isolation hypothesis. Examination of the overall recall scores, however, indicates that the groups were by no means matched in this respect. Random sampling errors seem to have placed the better memorizers in the change group and the worse in the control group. A similar difficulty was met in the preceding experiment and is dealt with in the same way. The result of this analysis is set out in Table 5.

This enables us to plot the probability of recalling item *4* as a function of overall recall scores for each of the three groups, as is shown in Fig. 5. Our hypothesis would lead us to predict that the graphs for the experimental groups will lie above that for the control group, since when overall recall ability is equated the probability of recalling item *4* should be greater under change and isolation conditions.

This general tendency is apparent, the graph for the control group only once going above either of the experimental groups. Using the same statistical treatments as were introduced in the first experiment we find that comparing the control and change groups

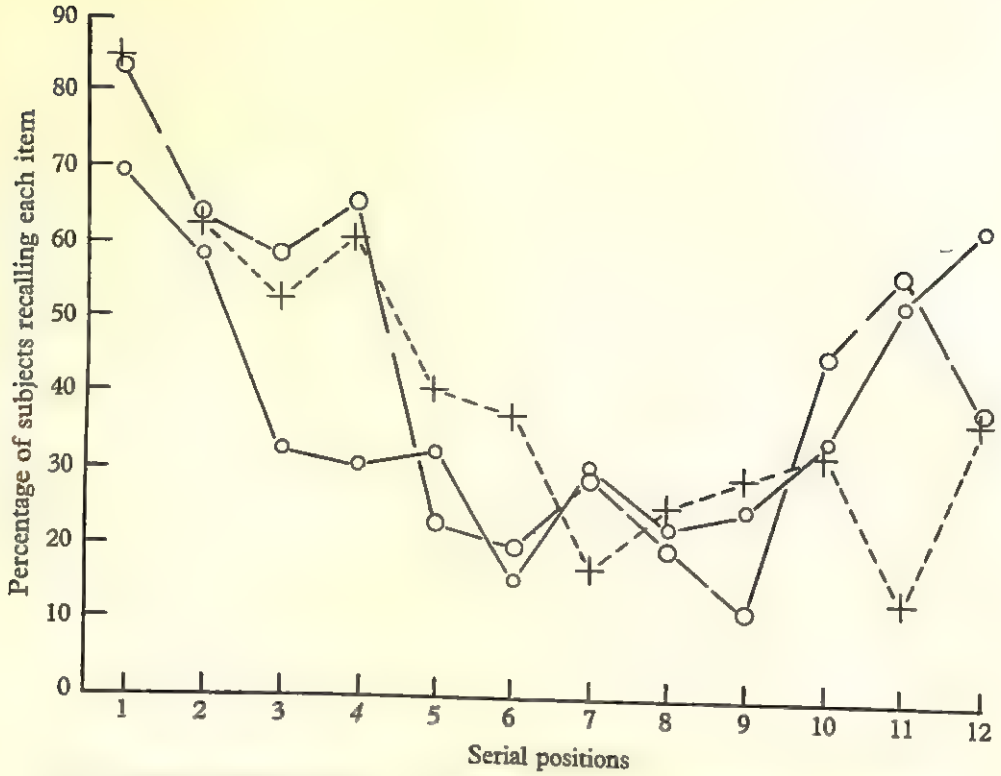


Fig. 4. Serial position curves. o——o, control; O– – –O, change; +- - -+, isolation.

Table 4. *Recall frequencies of item 4*

	Control group	Change group	Isolation group
No. of subjects	53	55	59
R4	16	36	36
F4	37	19	23
Recall (%)	30·2	65·5	61·0
Average no. items per subject	4·64	5·13	5·02

Control *v.* Isolation

$$\chi^2 = 8{\cdot}6^2\left(\frac{1}{27{\cdot}4}+\frac{1}{31{\cdot}6}+\frac{1}{24{\cdot}6}+\frac{1}{28{\cdot}4}\right)$$

$$= 10{\cdot}6 \qquad P \simeq 0{\cdot}001 \text{ (two-tail)}$$

R4, those subjects recalling item *4*; F4, those subjects forgetting item *4*.

Table 5. *Equating sub-groups for recall ability*

Total no. items recalled	Control group R4	Control group F4	Control group *P*	Change group R4	Change group F4	Change group *P*	Isolation group R4	Isolation group F4	Isolation group *P*
1	0	2	0·00	0	0	—	0	1	0·00
2	0	3	0·00	1	3	0·25	0	2	0·00
3	1	8	0·11	3	5	0·38	3	4	0·43
4	6	8	0·43	7	4	0·64	5	5	0·50
5	0	9	0·00	7	2	0·78	7	6	0·54
6	3	3	0·50	6	3	0·67	14	4	0·78
7	3	4	0·43	6	1	0·86	5	1	0·86
8	1	0	1·00	3	1	0·75	1	0	1·00
9	2	0	1·00	2	0	1·00	0	0	—
10	0	0	—	1	0	1·00	0	0	—
Total	16	37		36	19		36	23	

R4, those subjects recalling item *4*; F4, those subjects forgetting item *4*.

by the first method: $t=2{\cdot}32$; $\eta=7$; $P\simeq0{\cdot}06$ (two-tail). For the control and isolation groups: $t=2{\cdot}65$; $\eta=7$; $P\simeq0{\cdot}05$ (two-tail). There is no discernible difference between the change and isolation groups beyond a faint trend of no significance in favour of the change group. The exact treatment leads to closely similar conclusions.

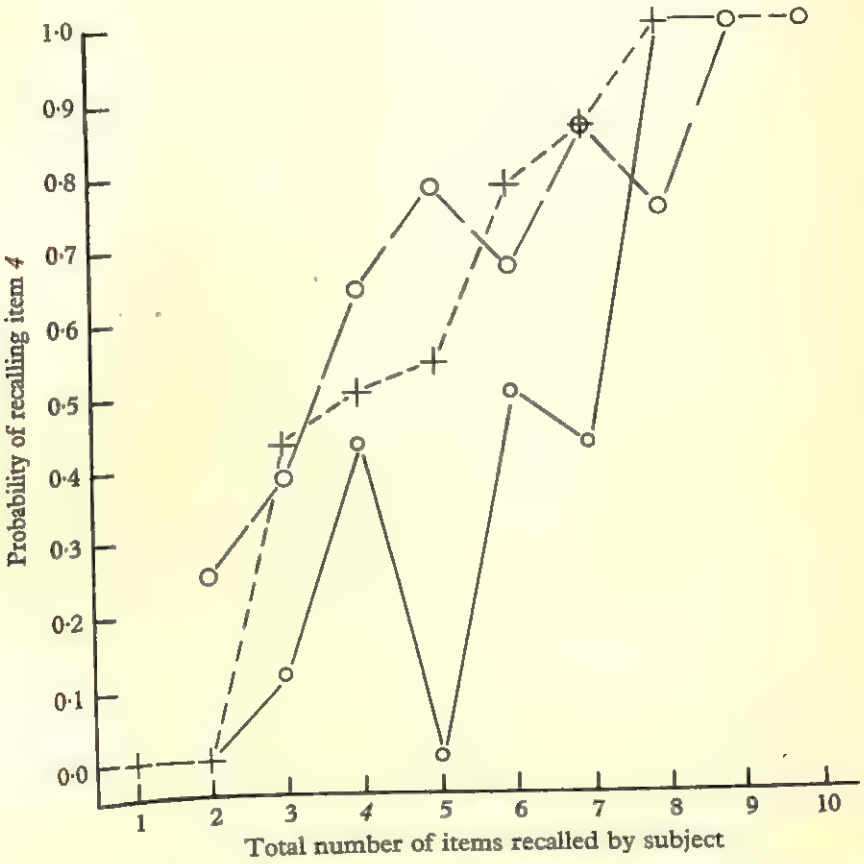


Fig. 5. Recall probability of item 4. o——o, control; O——O, change; + - - - +, isolation.

IV. *Discussion*

The obvious conclusion to be drawn from this result is that isolation in the von Restorff sense of the term has no effect. The relevant variable in this case appears to be structural change and only when an isolated item introduces such a change is it favoured in recall. So it would seem that earlier findings, including some of quite recent origin such as that of Kimble & Dufort (1955), which produced results favouring the isolation hypothesis, are an artefact resulting from the confounding of variables.

Another way of interpreting the data for *CZ* is in terms of pro-active interference. Structural change as defined in this experiment means introducing an item of a new type. So it could be argued that the main reason why the critical item is recalled less often by the control group is that *CZ* has been preceded by another letter item, *RQ*, which in some way either inhibits the perception of, or interferes with the engram of the critical item. For our present purposes, whatever the alternative explanation may be, the evidence is sufficient to discredit the isolation hypothesis.

References

Gibson, E. J. (1940). A systematic application of concepts of generalisation to verbal learning. *Psychol. Rev.* **47**, 196–229.

Green, R. T. (1956). Surprise as a factor in the von Restorff effect. *J. Exp. Psychol.* **52**, 340–4.

Kendall, M. G. (1948). *Rank Correlation Methods*, p. 35. London: Griffin.

Kimble, G. A. & Dufort, R. H. (1955). Meaningfulness and isolation as factors in verbal learning. *J. Exp. Psychol.* **50**, 361–8.

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychol. Forsch.* **18**, 299–342.

(*Manuscript received* 26 *November* 1956)

IMPOSSIBLE OBJECTS: A SPECIAL TYPE OF VISUAL ILLUSION

By L. S. PENROSE and R. PENROSE

(University College, London, and Bedford College, London)

Two-dimensional drawings can be made to convey the impression of three-dimensional objects. In certain circumstances this fact can be used to induce contradictory perceptual interpretations. Numerous ideas in this field have been exploited by Escher (1954). The present note deals with one special type of figure. Each individual part is acceptable as a representation of an object normally situated in three-dimensional space; and yet, owing to false connexions of the parts, acceptance of the whole figure on this basis leads to the illusory effect of an impossible structure. An elementary example is shown in Fig. 1. Here is a perspective drawing, each part of which is accepted as representing a three-dimensional rectangular structure. The lines in the drawing are, however, connected in such a manner as to produce an impossibility. As the eye pursues the lines of the figure, sudden changes in the interpretation of distance of the object from the observer are necessary. A more complicated structure, not drawn in perspective, is shown in Fig. 2. As this object is examined by following its surfaces, reappraisal has to be made very frequently.

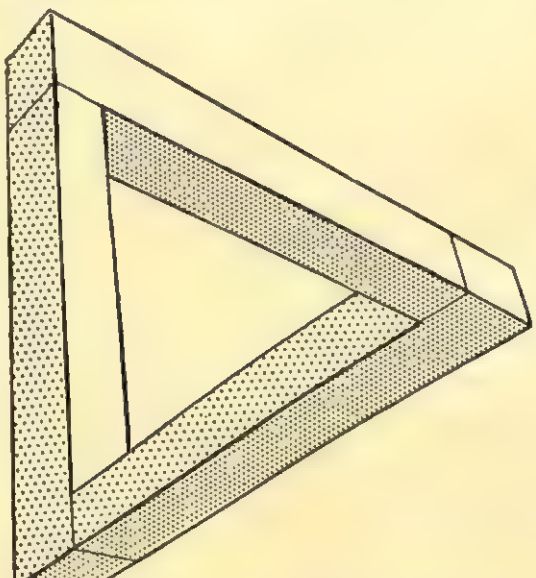


Fig. 1. Perspective drawing of impossible structure.

Another way of presenting the same type of illusion is to express the impossibility in terms of such a phenomenon as a continually descending or ascending path. The flight of steps drawn in Fig. 3 is an example of this. Each part of the structure is acceptable as representing a flight of steps but the connexions are such that the picture, as a whole, is inconsistent; the steps continually descend in a clockwise direction.

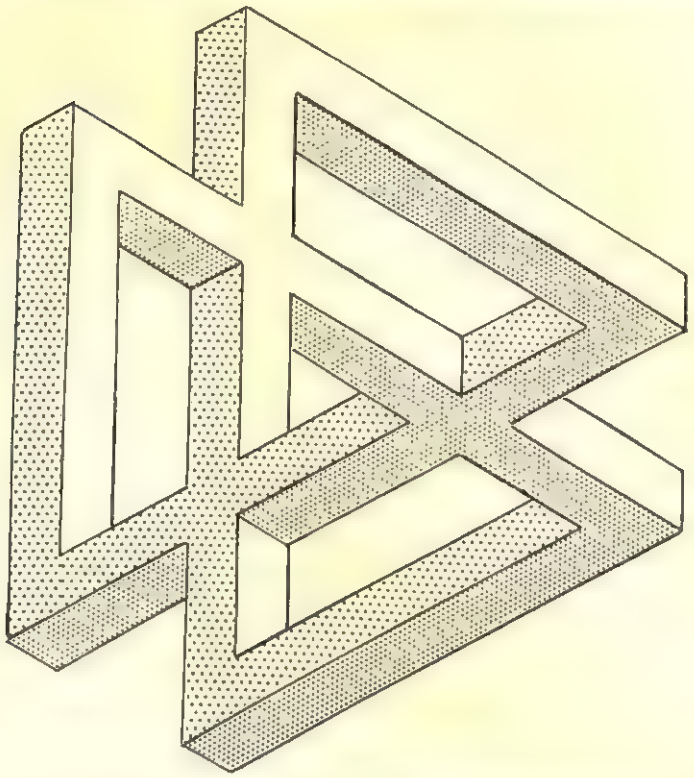


Fig. 2. Diagram of structure with multiple impossibilities.

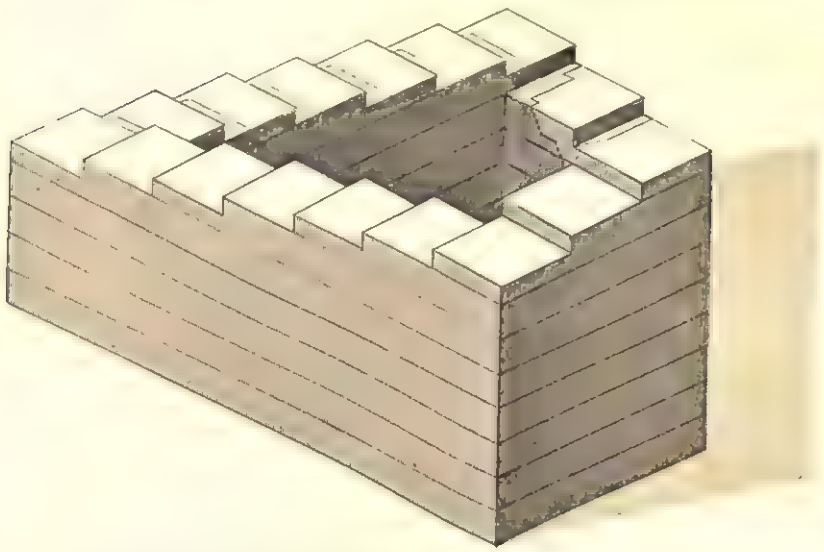


Fig. 3. Continuous flight of steps: shadowed drawing.

Actual objects suitably designed, when viewed from particular angles, can give exactly the same impressions as inconsistent drawings. A photograph of a model of this kind, apparently an impossible staircase, is shown as Fig. 4. Actually the extreme right hand step was much nearer to the camera and on a much higher level than the step which appears to be just above it. Illusions with a different purpose, constructed in a somewhat similar manner, have been discussed by Kilpatrick (1952).

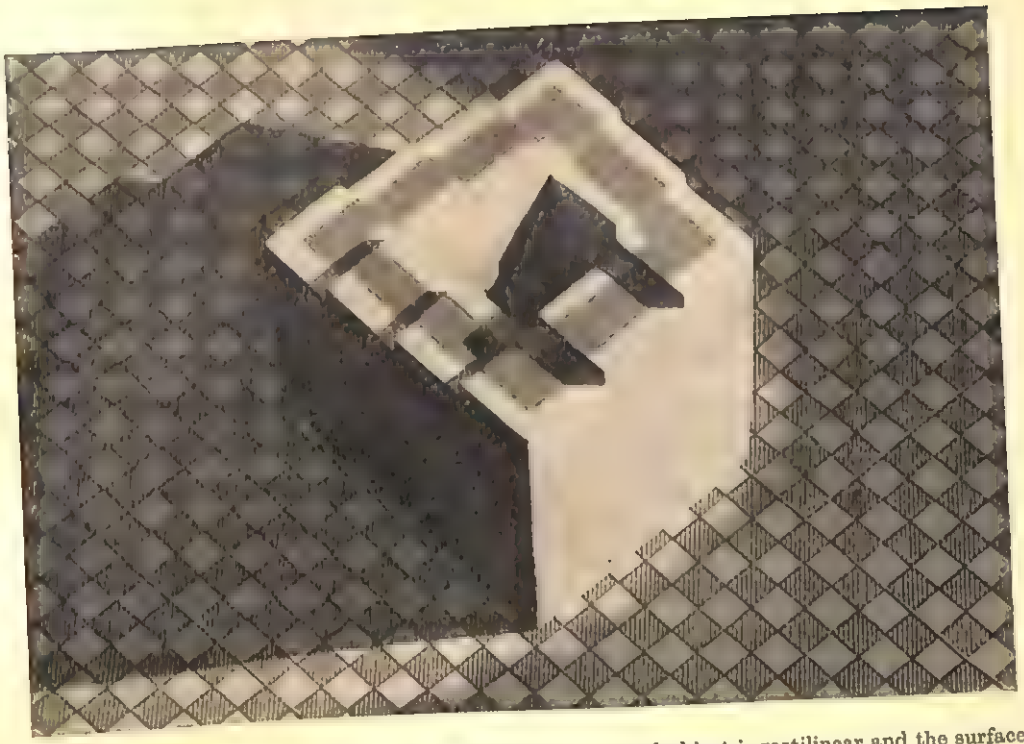


Fig. 4. Untouched photograph of continuous steps: the actual object is rectilinear and the surfaces of the steps are horizontal.

REFERENCES

ESCHER, M. C. (1954). *Catalogus* 118; Stedelijk Museum, Amsterdam.

KILPATRICK, F. P. (1952). Elementary demonstrations of perceptual phenomena. In *Human Behavior from the Transactional Point of View.* Hanover, N.H.: Institute for Associated Research.

(*Manuscript received* 30 *November* 1956)

THE EFFECTS OF THE MENSTRUAL CYCLE ON TEST PERFORMANCE

By MARY WICKHAM

It has been suggested that the menstrual cycle may represent a source of unreliability in the group testing of older girls and women. Data are presented for 4000 young women in which test scores are related to menstrual and health variables. It would appear that, during the Period Phase, there tends to be a slight depression of scores, other than in practical/mechanical performance, but this is too inconsistent and insignificant to warrant compensation in prediction, being no greater than the effects of the more common minor ailments.

When it is necessary to test women in large groups some of them will be passing through the physical disturbance of menstruation which may thus seem to present a problem additional to that of the usual minor ills. The aim of this investigation was to findout whether this disturbance influenced test performance, and if so to what extent. Two groups of women were used as subjects. Group A contained 3000 and group B 1000. Members of group A, all of whom had been in the A.T.S. for at least six months, were attending a Selection Centre with a view to changing their jobs. There they were interviewed, tested, and reclassified. The operation lasted from early February until late July so that weather conditions varied considerably. This group proved to be fairly representative of the service as a whole in terms of Progressive Matrices scores and educational background, but had more women of low medical category than was to be expected by chance (Table 1). Group B consisted of recruits who were first tested during a spell of extremely cold weather in February and March. They, too, were representative in terms

Table 1. *Distributions of both groups compared with a representative sample*

Selection Grade*	Group A (%)	Group B (%)	Representative sample (%)
(a) By general ability (on the Progressive Matrices)			
1 Top 10%	16·2	16·3	18·8
2 Next 20%	26·0	32·7	28·5
3+ Next 20%	23·6	24·3	23·0
3− Next 20%	16·5	16·3	15·0
4 Next 20%	15·0	10·4	12·8
5 Bottom 10%	2·7	—	1·1
(b) By educational background			
Left school after 16, with School Certificate (or better qualification)	5·5	6·5	5·7
Left school at 16 without School Certificate	14·0	11·5	11·9
Left school at 15 or attended night school	12·9	15·5	18·0
Left school at 14 from top class	55·2	53·5	52·2
Left school at 14 but not from top class	12·4	13·0	12·2
(c) By medical category†			
A.W.	74·5	95·6	89·5
B.W.	11·4	4·4	9·5
C.W. and E.W.	14·1	—	1·0

* See Vernon & Parry (1949), p. 177.

† The medical categories used for the classification of the women were very similar to those used for the men. Thus A.W. was the classification for those of good physical standard; B.W. for those with some physical fault or handicap (frequently a slight defect in eyesight might account for this); while in the C.W. and E.W. categories the faults were more marked and serious.

of test scores and educational background, but had an unduly large proportion in high medical categories (Table 1).

Both groups of subjects did the following tests, all of which are fully discussed by Vernon & Parry (1949):

Test	Items	Reliability
S.P. 1. *Progressive Matrices* (Raven, 1939)	60	0·88
S.P. 2. *Mechanical Comprehension* (Bennett)	50	0·85
S.P. 23. *Arithmetic*	30	0·89
S.P. 4. *Squares test.* This spatial test consisted of figures in which the testee had to draw a dividing line so that the two pieces formed would, if turned round, make a square	50	0·85
S.P. 14. *Spelling.* Five incorrect and one correct version of each word were given	50	0·91
S.P. 21. *Instruction test.* This test involved the comprehension of rather complicated instructions	150	0·92
S.P. 24. *Mec. test.* This was a practical test involving the assembly of Meccano parts to a model with the aid of diagrams	7	0·77
S.P. 25. *Verbal test.* This was based on synonyms and rhymes	104	0·97

Other information was collected under a number of heads, the procedure being different as between the two groups. The data listed below were available.

Group A

When these women were at the Selection Centre their medical category was checked; they were given the whole battery of tests described above; they were interviewed by a selection officer, a medical officer, and if, in the opinion of either of these officers, it was thought desirable, they were also interviewed by a psychiatrist. In the course of this procedure the following information was obtained for each woman: (i) scores on the whole test battery described above; (ii) re-checked medical category; (iii) educational background; (iv) service and previous civilian employment; (v) date of the onset of the immediately previous menstrual cycle; (vi) average length of the menstrual cycle; (vii) average duration of the menstrual period; (viii) incidence and severity of dysmenorrhoea; (ix) a rating on mental stability (this rating was given by the selection and medical officers and in cases where they suspected serious difficulties the opinion of the psychiatrist was included); (x) the original Matrices grade which had been obtained when the woman joined the service, taken at least six months earlier and at random in so far as the menstrual phase was concerned.

Group B

These women were tested with the whole test battery and then retested with the same battery in six weeks time. Each was interviewed and medically examined and herself filled in a questionnaire. During the course of this procedure the following information was obtained for each of these women: (i) scores on the whole test battery; (ii) retest scores for the same; (iii) medical category; (iv) educational background; (v) date of the immediately previous and of the next expected menstrual period (this was obtained at both the time of the test and retest); (vi) any feeling of dysmenorrhoea at either the test or retest; (vii) any other factor which they felt to be having a disturbing influence at either test or retest (the common cold and chilblains were the handicaps most frequently named, being a reflexion of the severity of the weather at the time of the investigation).

In previous work on this subject the investigations, such as those of Sowton & Myers (1928), had been concerned with the performance of individuals over a series of menstrual

cycles. A tendency to depression in functional activity shortly before and during menstruation and a heightening in the middle of the inter-menstrual phase, had been noted by such writers as Bedale (1928). The findings have been far from consistent. A general tendency for menstruation to depress test scores had been claimed. Some divergence between the effect in the performance of the same individual in different tests had also been noted by Sowton & Myers (1928), the scores in the tests of an intellectual type being depressed while those of manual skills were in some cases even increased. In most of this previous work the limitations and difficulties inherent in the tracing of the performance of individuals over many cycles were stressed. In this investigation some of these difficulties were avoided. A wide range of tests was used, involving about two hours testing time; such could obviously not be done in daily repetition. The tests were given on two occasions only, but to large numbers. The women tested did not think the information they gave about their menstrual cycle of much import, as it seemed merely incidental to the main aims at the testing.

The results in the second series of testing were compared with those of the first to see to what extent they varied beyond expectation, when the subjects were passing through some disturbed phase in the menstrual cycle.

The data obtained from group A were dealt with as follows. Length and duration of the menstrual cycle were considered. No less than 80% of this group gave the usual length of their menstrual cycle as 28 days, 8% gave a shorter and 12% a longer cycle than this, varying from 14 days on the one hand to 43 on the other. There was a marked tendency for the cycle to be considered in multiples of 7 and given in weeks rather than days, thus there were peaks rising to 30 individuals at 14 days, 120 at 21 days, 70 at 35 days and a minor peak at 42 days.

It is obviously difficult to assess the actual length of 'period' with accuracy, as even within the experience of one individual there are fluctuations. Some 94% of this group quoted their 'period' as lasting from 3 to 7 days, over 800 women giving 5 days as the usual duration. Of the remainder, 6%, some claimed a 'period' as short as 1 day, while there was one case in which it lasted 14 days.

Incidence of dysmenorrhoea

During the course of the interviews it was found that 24·4% of the group experienced some pain immediately before or during the 'period'. For the majority the pain was felt during the first or the second day. These figures are not very different from those of other writers; thus Sturgis (1923) found that 30·6% of over 2000 women employed in a store suffered some slight handicap while 4·4% found it serious. The incidence of dysmenorrhoea was greatest in the lower medical categories and in the more strenuous jobs, thus it was significantly more marked in the operational members in anti-aircraft than those occupied in the more sedentary jobs of clerking and signals. This is shown in Table 2.

Table 2. *Significance of differences in the percentage incidence of dysmenorrhoea*

(*a*) By *medical category:* A.W. 23·28%, B.W. 26·20%, C. and E.W. 29·64%
% difference: A.W. – C. and E.W. 6·36; standard error 2·78

(*b*) By *job:* operational nos. 27·0%, cooks 25·1%, clerks and signals 21·9%
% difference: operational nos. – clerks and signals 5·1, standard error 2·5

(*c*) By *normality of cycle and period:* normal 19·54%, abnormal 29·26%
% difference: normal – abnormal 9·72, standard error 1·99

Owing to the great variation in length of cycle (noted above) a subdivision was made into groups giving the *Normal Cycle* (28 days), a *Short Cycle* (20, 21 or 22 days) and a *Long Cycle* (34, 35 or 36 days).

Two outstanding phases within the menstrual cycle have been shown by such writers as Hoskins (1941) and Benedek & Rubenstein (1939). In mid-cycle, the build-up of oestrogen and progestin culminates in ovulation. Towards the end of the cycle, the destructive forces are in control and they culminate in the menstrual period. Since the stimulating build-up and the following break-up take place gradually there are two other intervening phases of 'rest'. From 4 days before to 4 days after the onset of menstruation would seem to cover the period when the destructive and depressing phase is of greatest influence, and from the 11th to the 18th day would cover the ovulation, when the stimulation would be at a maximum in those women who had the normal length of cycle, and the normal length of period. The mid-cycle 'ovulation' phase was of equal interest to that of menstruation as it might have had a stimulating effect on the test scores. The nomenclature used by Benedek & Rubenstein (1939) was adopted, namely the 'ovulation' and the 'menstrual' phase and the two intervening phases.

Bearing these complexities in mind, and in view of the relatively large numbers of women whose results were available, it was decided to minimize the chance variables by including in this subgroup only those quoting a normal length of cycle and period. There were 1335 for whom this was so and for whom all other relevant information was available. (In a few cases some of the earlier test results were not known and they had to be excluded. Most of the relevant data were available for 1525 cases.)

The 'short' and the 'long' cycle subgroups were similarly considered, the period phase being taken as from 3 days before until 3 days after the onset of menstruation, and from 5 days before to 5 days after the onset for the two subgroups respectively. There were seventy-nine cases claiming a 'short' cycle and fifty-five who claimed a 'long' cycle. As the numbers here were smaller than in the normal group, and there was little evidence for placing the 'ovulation' phase, the results when grouped for 'phases' were considered in relation only to 'period' and 'non-period' phases.

The cases whose results did not fall into one of these subgroups were ignored for the next part of the analysis. It was felt that more would be lost than gained by including all the intervening and outlying cases as the variables involved (such as abnormal length of cycle and duration of period) might well mask rather than elucidate any influence of the menstrual cycle upon the test results.

The 'normal' cycle subgroup of 1525 cases was then sorted into 'day' groups so that all the women who were at the second testing on the same day within the menstrual cycle were put into the appropriate 'day' group. This was done for each of the 28 days of the cycle. The mean of the 1525 cases was found for each of the tests given. The average deviation of each of the 'day' subgroups was found for each of the tests. Thus the *actual deviation* for each day within the menstrual cycle was obtained for the eight tests described on page 35.

The results of the Matrices test, which each of these women had taken at least six months earlier, had been scored at random, in so far as menstruation was concerned. The average score of each of the 'day' subgroups was again found, and then the deviation of each 'day' subgroup from the average of the 1525 (forming the whole group) having a normal cycle. Thus the expected deviations were obtained. For tests other than the

Matrices, the method given by G. W. Snedecor (1940) was followed. The inter-correlation of each of the tests with Matrices was obtained and the reliability of all the tests was known (see page 35). Thus the expected deviations were obtained.

The results of those falling into the phases as regards the second testing (i.e. the Period, the Ovulation, and the two intervening phases of rest as described on page 37) were combined to give the observed deviations for the 'phase' subgroups. These were then considered against the expected deviations.

If the different phases affect performances on these tests differently, we can expect that the average performances of different phase groups will vary. On the other hand, if we find that average performances vary, we cannot infer that this is due to differences in phase, since the normal standard of performance of the individuals in each phase group may be different. There are, thus, two possible explanations of any observed differences between the phase groups. Knowing the grades on the original testing with Progressive Matrices, which are independent of phase, we can estimate how far the phase groups differ in normal standard of performance. We can thus establish expected differences between

Table 3. *Significance of the excess of differences between the observed and expected performances*

Test	*F* value	*P*
S.P. 2. Mechanical Comprehension	1·7	>0·05
S.P. 23. Arithmetic	2·08	>0·05
S.P. 4. Squares	1·33	>0·05
S.P. 14. Spelling	1·65	>0·05
S.P. 21. Instructions	3·80	<0·05 >0·01
S.P. 24. 'Mec' test. Practical models	2·39	>0·05
S.P. 25. Verbal	1·32	>0·05
S.P. tests (sum of scores for all tests)	1·04	>0·05

the phase groups and measure how far the observed differences exceed the expected differences. The excess can be attributed to the effects on performance of different phases in the cycle. The error variance of the measurements can be derived from similar measurements comparing observations with expectations for separate day groups within the same phase group, and an *F* test can be applied to measure the significance of any excess found.

This preliminary review was made to see which of the tests warranted further analysis. The *F* values obtained for the significance of the variance 'within' as compared with the variance 'between' the four phases are given in Table 3.

S.P. test 1, the Matrices, is omitted from Table 3 as it was decided to make a more detailed analysis of this test. From the results it was concluded that the Instructions test also warranted further analysis. The test results of the women who were in the 'short' and 'long' cycle subgroups were analysed in the same way. Although slightly lower scores than expected were obtained on most of the tests by the women in the 'Period' phase, in no case was significance reached, it being most nearly approached in the Instructions test.

At this point it seems to be worth noting that the incidence of dysmenorrhoea was greatest in those who had abnormal cycles (see Table 2). For the group as a whole (all 3000), those claiming a normal cycle and period were compared with those falling outside these limits in either direction for either cycle or period in relation to the incidence of dysmenorrhoea. Since a very significant difference is found, the test results of the abnormal groups might be expected to show the greater disturbance in score.

There is a tendency for a depression of score on the paper and pencil tests to occur for those in the 'period' phase. This tendency is reversed in all groups, for the practical Mec. test, and in the Arithmetic and Instructions tests for the group with an abnormally long cycle.

After this preliminary work, a more exhaustive analysis was undertaken for the Matrices and the Instructions tests. The Matrices was chosen because the initial test grades were available. The test-retest correlation was 0·790. The Instructions was chosen because, in the preliminary work, the difference in score had most nearly reached significance (see Table 3).

Table 4. *Analysis of variance of the 'days' within the menstrual cycle (Matrices test)*

Source of variance	Sum of squares	Degrees of freedom	Mean square
Total variance on retesting	1194·244	1334	—
Accounted for by regression of retest on test	744·350	1	—
Accounted for by differences in performance on different days	13·396	27	0·496
Residual variance	436·498	1306	0·334

$F = 1{\cdot}48$, $P > 0{\cdot}05$.

Table 5. *Analysis of variance due to menstrual phase (Matrices test)*

Source of variance	Sum of squares	Degrees of freedom	Mean square
Accounted for by differences between			
(*a*) Phases	0·7878	3	0·263*
(*b*) Days in the same phase	12·609	24	0·525†

* $F = 1{\cdot}27$, $P > 0{\cdot}05$. † $F = 1{\cdot}57$, $P < 0{\cdot}05 > 0{\cdot}01$.

Table 6

Source of variance	Sum of squares	Degrees of freedom	Mean square
Total variance on retesting	5247·1174	1524	—
Accounted for by regression	2085·8822	1	—
(i) Of Instructions on Matrices	10·0354	3	3·35
(ii) Differences within phases	35·5885	24	1·48
(iii) Differences between days within the same phase	3115·6113	1496	2·08
Residual variance			

F (ii) 1·61, $P > 0{\cdot}05$.
F(iii) 1·40, $P > 0{\cdot}05$.

The differences between the observed and the expected scores of the women whose retest fell on each of the 28 days of the menstrual cycle were first obtained separately and analysed (see Table 4). In spite of the large numbers involved the differences are only on the borderline of significance. The analysis was then carried a step further. All the women tested during the same 'phase' were grouped together (see Table 5). This table shows that such differences as occur between observed and expected mean retest scores are irregular and not *significantly* associated with a *particular day or phase* within the menstrual cycle.

A similar analysis was then made for the Instructions test. The mean scores were estimated from the results of the original testing on the Matrices. The correlation between the Matrices and Instructions tests for this sample was 0·631 (see Table 6).

The significance indicated for the Instructions test in the preliminary analysis is not confirmed.

It can then be concluded from the results of this group of women that no advantage would be gained by taking into account the day within the menstrual cycle on which the test was given for predictive purposes.

In group B the data were dealt with as follows. Since the test and retest results for each test of the battery were available for each of the 1000 women in this group, a simple method was available. There was a 6-week interval between the first and second testing, and so the results obtained for each woman were for some different day within the menstrual cycle. If the menstrual phase influenced scores adversely, then the women in this phase at the first testing would not, at that time, do themselves justice. At the retest these women, not being in this phase of expected depression, would show their real capacity. Therefore the differences between their test and retest scores would be *greater* than the differences between the test scores of the women who were not in the menstrual (period) phase at either testing or retesting. Conversely, the women in the menstrual (period) phase at the time of retest would be handicapped at the second testing, and the difference between their test and retest scores would be less than those of the women who were not in this phase at either time.

Table 7. *The means of the differences in test scores*

Test	$(b-a)$ expected positive	S.E.	$(c-a)$ expected negative	S.E.
S.P. 2. Mechanical, paper and pencil	+0·109	0·642	1·440	0·814
S.P. 23. Arithmetic	0·943	0·773	−0·559	1·106
S.P. 4. Squares	−0·970	0·665	1·176	0·730
S.P. 14. Spelling	−0·743	0·574	0·075	1·025
S.P. 21. Instructions	−0·139	1·228	−0·377	2·070
S.P. 24. Mechanical, practical	−3·55	1·50	−2·325	2·00
S.P. 25. Verbal	0·760	0·733	0·191	0·909

In order to examine these two hypotheses the means of the differences between the test and retest scores were obtained for: (*a*) those outside the menstrual phase at both test and retest; (*b*) those in the menstrual phase at the initial testing; and (*c*) those in the menstrual phase at retesting. The differences of the means for the above groups are given in Table 7. In view of the above hypotheses the value for (*b*) is expected to be greater than that for (*a*), and that for (*c*) is expected to be less than that for (*a*).

As can be seen, all differences are insignificant, except in one case (S.P. 24), where the difference is in the direction contrary to that expected.

As was noted on page 35, the common cold and chilblains were the most frequently quoted cause of disturbance in this group other than dysmenorrhoea. At the time of retesting no fewer than 177 reported severe colds.

In order that the effect of these minor ills could be compared with that of dysmenorrhoea, parallel calculations were made for the results of the Instructions and the practical Mechanical tests, as in the above work they had shown the most nearly significant results (see Table 8). Again the values found give no consistently significant trend. It is of interest to note that here there is no tendency for the Mechanical test to behave differently from the Instructions test as there was in the comparison where dysmenorrhoea was the handicap under consideration.

Table 8

	Instructions	Mechanical
Showing the differences between the means of the scores of those:		
Handicapped by cold at the test—The not-handicapped	−0·765	0·550 (positive value expected)
S.E.	2·05	2·09
Handicapped in the retest—The not-handicapped	−0·882	−1·303 (negative value expected)
S.E.	0·818	0·952
Handicapped by chilblains at retest The not-so-handicapped at the retest	−1·794	−1·269 (negative value expected)
S.E.	1·462	1·551

CONCLUSIONS

(1) In the results of the 3000 women in group A there was a tendency for depression of score in the Period Phase. This was noted for just over 70 % of the performances of the subgroups for the eight tests considered. (2) Half the exceptions to this tendency occurred in the results for the practical Mechanical test where the reverse was found, the women doing better during the Period Phase. (3) The depression of scores did not however reach a level of consistent significance. (4) The results of the 1000 women in group B were marked by inconsistency and a lack of significance. The influence of minor common ills such as colds and chilblains seemed equally likely to cause a fall in score. (5) The scores in the practical Mechanical test tended to be higher during the Period Phase, as for the first group. The scores in this test when the testees were under the handicap of colds, etc., were lower, behaving in the same way as the scores in the other tests.

Finally, it was concluded that no practical advantage would be gained by taking into account the day within the menstrual cycle on which a test was given for predictive purposes.

It will be appreciated that in an investigation of this kind the writer is grateful to the many officers and other ranks who gave and collected the data from which the conclusions were drawn, and above all to Dr Edith Mercer who made the investigation possible, and to Mr Patrick Slater and Prof. P. E. Vernon for their help and advice upon the statistical methods used.

REFERENCES

BEDALE, E. M. (1928). Experimental study of the menstrual cycle. Its relation to general functional activity. *Rep. Med. Res. Coun. Industr. Res. Bd*, no. 45.

BENEDEK, T. & RUBENSTEIN, B. B. The correlations between ovarian activity and psychodynamic processes. *Psychosom. Med.*, **1**, 244, 461.

Bodily disturbances during the menstrual period. In *Practitioners' Encyclopedia of Midwifery and the Diseases of Women*. London: Frowde and Hodder and Stoughton. 1921.

HOSKINS, R. G. (1941). *The Glands and their Functions*. London: Routledge & Kegan Paul.

SNEDECOR, G. W. (1940). *Statistical Methods*, chapter 12. Iowa State College Press.

SOWTON, S. & MYERS, C. S. (1928). Experimental study of the menstrual cycle. Its influence on mental and muscular efficiency. *Rep. Med. Res. Coun. Industr. Res. Bd*, no. 45.

STURGIS, M. C. (1923). Observations on dysmenorrhoea occurring in women employed in a large departmental store. *J. Industr. Hyg.* **5**, 53.

VERNON, P. E. & PARRY, J. B. (1949). *Personnel Selection in the British Forces*, p. 177. London University Press.

(*Manuscript received* 4 *December* 1956)

A 'CRESPI EFFECT' IN MALE IMBECILES

By N. O'CONNOR and G. S. CLARIDGE

Medical Research Council, Social Psychiatry Research Unit, Institute of Psychiatry, London

In order to test the effects on performance of an interchange in incentive level forty male imbeciles were divided into four equal matched groups on the basis of two initial trials on a simple dexterity task. For a further ten trials two groups worked under non-incentive conditions and two under incentive conditions, in which they were encouraged to reach a goal based on their previous performance. On the last trial one incentive group was changed to control conditions and one control group to incentive conditions, the other two groups continuing as before. The experiment was continued for a further seven trials.

Results indicated that groups given the incentive from the beginning of learning separated significantly from those receiving no incentive. When one of the latter groups was given an incentive, however, its performance increase showed a typical Crespi type 'elation' effect, performance reaching, and maintaining, a level greater than that of subjects who had received the incentive from the beginning of learning. The effect was found to be a group trend rather than a response typical of all subjects.

A comparable decrease in incentive level was found to have little marked effect on performance level, which established itself at a point intermediate between the performance levels of subjects who continued to receive an incentive and of subjects who had received no incentive at all. This finding confirmed earlier work with defectives, showing that incentives would have relatively permanent effects on the attitude or set, and hence on the performance level of imbeciles.

I. Introduction

It is generally agreed by learning theorists in both the animal and the human fields that changes in level of motivation will normally result in corresponding changes in performance level. The precise nature and extent of these performance changes have, however, been the subject of some controversy, particularly among comparative psychologists.

In the final treatment of his theory Hull (1951) gave specific consideration to this problem and derived a theorem to cover the effects of changes in the incentive component (K). He says:

> Other things constant, an abrupt shift in the incentive used during a maze-learning process will be followed first by a major shift in reaction potential and then by two or more progressively smaller shifts on successive trials, the series constituting a rapid learning process of the exponential variety, culminating in the course that the ${}_sE_R$ would have followed had the new incentive been operating continuously from the beginning of learning.

In deriving this theorem Hull (1951) makes use, somewhat paradoxically, of the work of Crespi (1942), who reports a very meticulous and elaborate study in which interchanges in a number of food amounts were made, the drive variable being carefully controlled at all stages of the experiment. Using speed of running as a measure of response, Crespi demonstrates changes in performance level which are greater than can be predicted from Hull's theorem. Thus, it was found that when a larger amount of food was introduced after previous training under a smaller amount, performance level increased to a point *above* that of animals who had received the larger incentive amount from the beginning of training. Crespi calls this excess increase an 'elation' effect and also records a corresponding 'depression' effect in animals shifted from larger to smaller amounts of food.

Similar contrast effects are also reported by Zeaman (1949) using latency as a measure of response.

While Hull explained what he himself termed the 'Crespi effect' as being merely a temporary, emotional phenomenon, Crespi (1944) attaches more significance to it. He suggests that the 'elation' effect, for example, is due to an increase in emotional drive as a result of the discrepancy between expectation, built up as a result of experience of a smaller incentive, and the sudden achievement of a larger incentive.

Much of the work carried out on changes in motivation in human subjects, which has tended to proceed independently of similar work in the comparative field, has been carried out with knowledge of results as the incentive situation. Mace (1935) and Book & Norvelle (1922) found rises and drops in performance when subjects were changed from no knowledge to knowledge of results and vice versa, but the results of Mackworth (1950) are not entirely consistent with these findings. He found that the introduction of knowledge of results produced a significant fall in the error score of radar watchers, but that the reverse procedure had no significant effect on performance.

In studies with humans differences in the kind of incentive and experimental task employed by investigators, as well as the failure to provide adequate control groups, have led to some uncertainty about the precise effects on performance of changes in incentive level, although on the whole. as pointed out by Ammons (1954), results have tended to corroborate the predictions of Hullian theory, the 'Crespi effect' not having been demonstrated in human subjects.

II. The experiment

(i) *Aims of experiment.* Previous work with imbeciles (Gordon, 1953; Claridge, 1956) had suggested that under certain conditions of motivation the 'elation' effect reported by Crespi (1942) might occur. It might be expected, for instance, that under conditions where subjects had been made accustomed over a relatively long period to performance on a monotonous routine task without incentive, a change in these conditions would produce the kind of 'emotional boost' proposed by Crespi (1944).

The outcome of a change in the reverse direction—that is, from more to less favourable incentive conditions—was, as in the case of normal human subjects, less easily determined. It was possible, in fact, that little change in performance would occur, for results reported by O'Connor & Claridge (1955) suggested that favourable incentive conditions early in practice would have effects on the performance of imbeciles which were persistent for periods of at least a year.

In order to throw some light upon these problems an experiment was designed which would provide for systematic increases and decreases in the level of motivation of imbeciles, with adequate control groups for these changes. It was hoped, thereby, to make a more precise comparison than has hitherto been possible with the current approach of comparative learning theorists to this problem.

(ii) *Subjects.* Subjects were selected from the imbecile population of the small branch institution of a large hospital for mental defectives. Forty patients between the ages of 18 and 45 were finally chosen, excluding those who were blind, epileptic, or severely crippled. Estimates of I.Q., made by means of the 'Draw a Man' test and scored according to the instructions of Goodenough (1926), confirmed the diagnosis of imbecile in all cases. Mean I.Q. for the 40 subjects was 41·02, S.D. 7·55 and mean age was 29·20, S.D. 7·71.

(iii) *The task.* The task used was a routine motor task termed the Nail Frame and devised by Gordon (1953). It consisted of a frame ($8 \times 10 \times \frac{3}{4}$) covered with a sheet of perforated zinc containing 82 holes to the square inch. A column down the shorter side of the board consisted of exactly 60 holes and subjects were asked to insert small escutcheon pins ($\frac{7}{16}$ in. in length) into the holes, working vertically down a column and completing it before starting the next. Each subject was given a small tray, containing a large supply of pins, which was placed at the side of the board appropriate to the handedness of the subject. Subjects were required to work for 60 minutes each day, this constituting a trial. The number of columns and the number of pins inserted in a partially completed column were easily calculated at the end of a trial by means of a graduated scale along the edges of the frame.

(iv) *The incentive.* The incentive condition used for the experiment was termed Goal with Encouragement, which had been shown (O'Connor & Claridge, 1955) to be maximally effective with male imbeciles. Under this incentive subjects were encouraged at intervals to reach a goal based on their previous performance. Achievement of the goal was praised and failure sympathetically observed by the experimenter.

(v) *Design and procedure.* Four groups of subjects were required and, since they were to be matched, two initial trials were given to form a basis for matching. On two successive days the 40 subjects were taken in groups of 10 and given an hour's trial on the Nail Frame, subjects simply being told to work as hard as they could. Two scores were, therefore, obtained for each subject, and on the basis of the mean of these scores subjects were assigned to one of four groups. The 40 subjects were ranked in order of initial performance level and split off into quadruplets, a member of each quadruplet being allotted at random to one of the four groups, which were designated as follows:

Goal-Goal (GG)	10 subjects
Goal-Control (GC)	10 subjects
Control-Control (CC)	10 subjects
Control-Goal (CG)	10 subjects

For the next 10 days (trials) of the experiment two groups (Goal-Goal and Goal-Control) worked under incentive conditions, which were as follows. *Goal with Encouragement conditions.* At the beginning of each trial a red nail was inserted in the board of each subject at a point which represented a 1% increase on the highest of his three previous scores. At the beginning of each trial subjects were told; 'Let's see you all get your goals today. You worked very well yesterday, but see if you can work even harder today.' Every quarter of an hour, they were told: 'Come on, boys, keep working hard. You are doing very well at the moment. Let's see you get your goals.' In addition to the group encouragement individual encouragement was given to each subject during the course of the trial. At the end of the trial those who succeeded in reaching their goals were told: 'That was very good; see if you can get your goal again tomorrow.' Those who failed were told: 'Bad luck; you did very well, anyway, and you'll get your goal tomorrow.'

For the first part of the experiment the other two groups (Control-Control and Control-Goal) worked under the following conditions. *Control or non-incentive conditions.* At the beginning of each trial subjects were told, 'work as hard as you can'; and at the end, 'you worked very well today; I'll see you all again tomorrow at the same time'.

At the end of this period, i.e. on the 13th trial of the experiment, a change in incentive conditions was made. One incentive group (Goal-Control) was changed to non-incentive conditions, the other incentive group continuing as before. At the same time one non-incentive group (Control-Goal) was changed to conditions of Goal with Encouragement, the other non-incentive group remaining on Control conditions. The experiment was then run for a further seven trials, at which point it was terminated.

III. Results

In order to convert the data into a form which would give measures of relative improvement and performance level, a linear regression technique was applied to the results, treating the pre-change period (trials 3–12) and the post-change period (trials 13–20) separately. The regression coefficients for linear regression (score $= a$(trials) $+ b$) for each individual are shown in Table 1, while Fig. 1 shows the group mean scores for each trial plotted as a graph.

The treatment of results in the periods prior to and following the change in incentive level will be considered separately in turn.

The pre-change period. It is apparent from Fig. 1 that at the point where the change in incentive was made the two incentive groups were clearly differentiated from the two non-incentive groups. This was found to be due, not so much to an increase in the over-all rate of improvement during this period, as to an abrupt increase in performance level when the incentive was introduced at the beginning of the experiment. Once the Goal-Goal and Goal-Control groups were established at this level the rates of improvement for

all four groups were closely similar, as shown by the following group means for the '*a*' coefficients during this period: Goal-Goal, 8·09; Goal-Control, 7·95; Control-Control, 7·27; Control-Goal, 7·60.

An Analysis of Variance confirmed that there was, in fact, no significant difference between the four groups in rates of improvement, F being 0·112. The superiority in level gained by Goal subjects on the first trial under these conditions was thus maintained throughout the pre-change period as revealed in the marked separation of the two pairs of groups. An Analysis of Variance computed in individual performance levels for the 12th trial indicated that there was a significant difference between incentive conditions at the end of the pre-change period ($F = 5{\cdot}965$, significant at 0·05 level). Although the significance level was a relatively lenient one, the clear over-all group trends and the way

Table 1. *Regression coefficients ('a' and 'b') for each subject in each group.*

		Trials 3–12		Trials 13–20	
Group	Subject	'a'	'b'	'a'	'b'
Control-Goal	1	9·9	738	−15·7	994
	2	20·0	612	−8·0	837
	3	0·8	744	4·7	981
	4	14·8	516	0·3	724
	5	3·7	596	1·0	767
	6	−19·6	674	7·1	668
	7	−2·2	346	−1·3	512
	8	27·3	417	5·5	662
	9	7·1	351	12·5	520
	10	14·2	326	5·9	528
		Mean 7·60	532·0	1·20	719·3
Control-Control	11	−3·6	862	4·7	851
	12	9·1	661	−5·0	833
	13	14·2	669	2·3	810
	14	6·6	580	−11·9	684
	15	−0·4	514	−10·4	482
	16	7·8	466	6·3	496
	17	16·5	424	−1·6	561
	18	28·1	357	−5·8	577
	19	−0·4	284	7·0	234
	20	−5·2	453	2·3	424
		Mean 7·27	527·0	−1·21	595·2
Goal-Goal	21	28·5	923	2·8	1138
	22	5·6	639	4·8	639
	23	4·4	672	−1·3	747
	24	11·8	623	1·0	824
	25	3·5	712	−9·6	845
	26	4·3	421	−11·2	500
	27	−4·1	430	−10·3	401
	28	24·2	480	−8·7	740
	29	−6·0	545	32·1	426
	30	8·7	420	10·9	486
		Mean 8·09	586·5	1·05	674·6
Goal-Control	31	11·1	778	−4·5	896
	32	−15·3	939	−6·3	815
	33	15·0	724	−4·9	865
	34	4·2	564	5·6	574
	35	5·9	619	11·6	644
	36	9·2	622	−4·8	688
	37	16·4	430	−2·3	574
	38	2·5	445	−2·9	483
	39	10·3	499	27·7	499
	40	20·2	214	−0·5	372
		Mean 7·95	583·4	1·87	641·0

in which the two Goal groups had kept together in separating from the two Control groups suggested that the differences were not due to chance but to a real effect of Goal with Encouragement.

The situation with regard to performance level, of the two pairs of groups at the point immediately prior to the interchange in incentive conditions indicated, therefore, that a critical test of the effect of this change could be made.

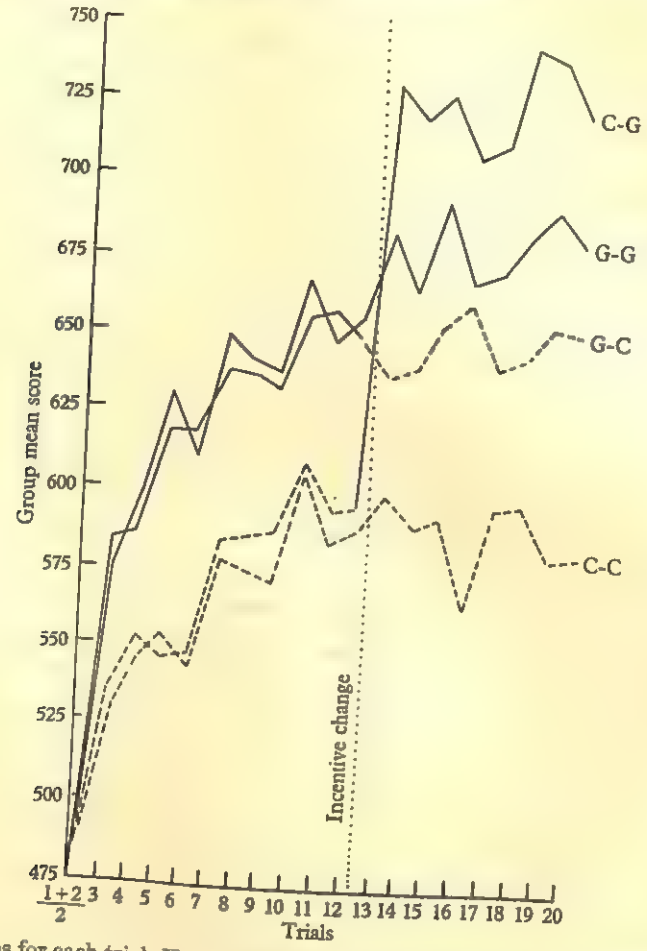


Fig. 1. Group mean scores for each trial. Key, incentive conditions: ——, goal; – – – –, control. Groups: G-G, goal unchanged; G-C, goal changed to control; C-C, control unchanged; C-G, control changed to goal.

The post-change period. The introduction of Goal with Encouragement conditions into the Control-Group on the 13th trial produced, as can be seen in Fig. 1, a marked increase in performance level, the group mean score on this trial being in the region of 125 points higher than that of the previous trial under non-incentive conditions. A *t*-test revealed that this increase was a significant one ('*t*' = 3·827 (D.F. 9) significant at the 0·01 level) and unparalleled in the other three groups. The level thus achieved was maintained throughout the post-change period, improvement being slight but positive. The average increment per trial for the group, as represented by the mean for the regression coefficient '*a*', was 1·20.

The mean performance level for the Control-Goal group did, in fact, reach a point which was above that of the Goal-Goal group on the 13th trial and remained consistently above

it for the remainder of the experiment. The 't' for the difference in group means for these two groups over the eight trials of the post-change period had a value of 14·053 (D.F. 7) which was significant beyond the 0·001 level. Inspection of individual scores indicated, however, that this 'elation' effect obtained for only some individuals, 't' calculated on the individual 'b' coefficients for the two groups being non-significant.

There was little marked effect of discontinuing Goal with Encouragement in the Goal-Control group, although some slight separation from the Goal-Goal group did occur. The difference in group mean score for this period between the Goal-Goal and Goal-Control groups was significant at the 0·001 level, 't' being 6·723 (D.F. 7), but the mean level of Goal-Control subjects was in turn significantly higher than that of Control-Control subjects. The value of 't' in this case was 8·554 (D.F. 7), significant at the 0·001 level. Inspection of individual scores revealed, however, that some subjects did decline following the removal of the incentive. Even in these cases, however, the decline was gradual rather than abrupt as in the case of the change in the reverse direction.

IV. Discussion and conclusions

Considered in terms of group trends it is apparent that the results reported here do not entirely bear out the predictions of Hull (1951). Although not clearly confirmed, because of individual differences, the marked reaction of imbeciles to the sudden introduction of an incentive supports, rather, Crespi's (1942) demonstration of a similar 'elation' effect. Broadhurst (1958) has recently confirmed some aspects of Crespi's work. It seems likely that this effect was to a large extent, as Crespi himself points out, a function of the level of expectation built up during the days when no reward was received. It must be remembered that the task was a highly monotonous and fatiguing one, even for imbeciles. By the change point, in fact, some dissatisfaction was being expressed by subjects. In the Control-Control group this became more marked from this point onwards and is reflected in the general decline of this group's performance level.

Two further points are of interest. First, it is noticeable that the 'elation' effect occurred in one trial, suggesting that, although they had previously performed at a lower level than subjects who had received an incentive during learning, individuals in the Control-Goal group had presumably acquired the same degree of skill on the task and were capable, given an incentive, of performing immediately at the level associated with that incentive. It is possible, in fact, as pointed out by Deese & Carpenter (1951), that the rate at which acquisition or habit growth occurs is independent of the level of drive. In any event, it confirms the necessity, suggested by many other authors, of distinguishing 'acquisition' from 'performance' in evaluating the role of motivation in learning.

The second point of interest is the relative permanence of the 'elation' effect. It is possible that, had the experiment been carried further, the performance level of Control-Goal subjects would eventually have trailed back to the level of Goal-Goal subjects, as stated by Hull (1951). There was, however, no evidence of this by the end of the experiment and it would probably have needed a great many trials for this to have occurred. It certainly could not be concluded that the effect was of a temporary nature.

The effect of removing incentives confirmed the findings of neither Hull nor Crespi and did not parallel the change in the reverse direction. It confirms, however, an earlier finding by O'Connor & Claridge (1955) that incentives will have a relatively permanent

effect on the performance level of imbeciles. The persistence of the attitude or set established under Goal with Encouragement conditions was illustrated by the practice of many subjects on the Goal-Control group of continuing to set themselves goals and show obvious pleasure when they claimed to have completed the number of rows they had set as their targets. The ability of imbeciles to 'learn' to set themselves goals in this way indicates that many of the present formulations of such concepts as habit, drive, and incentive may, by themselves, be too simple to explain the complex behaviour which is evident even at this level. This conclusion finds confirmation in many current approaches to the problem of motivation, such as those of Hebb (1949) and Harlow (1953).

At the same time the demonstrations that the 'elation' effect may be paralleled in human subjects suggest that it is through modification rather than rejection of current concepts that a more adequate theory of learning and motivation will emerge. The testing, with human subjects, of some of the basic postulates of animal learning theory would certainly seem to be a fruitful approach to this problem.

The authors wish to acknowledge their indebtedness to Dr R. J. Stanley, Physician Superintendent of Monyhull Hall, Birmingham, who made the patients available, and Dr H. C. Gunzburg, who selected the patients. The work reported was carried out by G. S. Claridge in partial fulfilment of the Ph.D. degree at the University of London.

References

Ammons, R. B. (1954). Knowledge of performance. Survey of literature, some possible applications, and suggested experimentation. W.A.D.C. Technical Report 54–14.

Book, W. F. & Norvelle, L. (1922). The will to learn: an experimental study of incentives in learning. *Pedagog. Sem.* **29**, 305–62.

Broadhurst, P. L. (1958). A 'Crespi Effect' in the analysis of emotionality as a drive in rats. *Brit. J. Psychol.* **49**, 56–8.

Claridge, G. S. (1956). Factors affecting the motivation and performance of imbeciles. Ph.D. Thesis, University of London Library.

Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.* **55**, 467–517.

Crespi, L. P. (1944). Amount of reinforcement and level of performance. *Psychol. Rev.* **51**, 341–57.

Deese, J. & Carpenter, J. A. (1951). Drive level and reinforcement. *J. Exp. Psychol.* **42**, 236–8.

Goodenough, F. L. (1926). *Measurement of Intelligence by Drawings.* New York: World Book Co.

Gordon, S. (1953). Some effects of incentives on the behaviour of imbeciles. Ph.D. Thesis, University of London.

Harlow, H. F. (1953). Motivation as a factor in the acquisition of new responses. In: *Current Theory and Research in Motivation: a Symposium.* Lincoln, Nebr.: Nebraska University Press.

Hebb, D. O. (1949). *The Organization of Behavior. A Neuropsychological Theory.* New York: Wiley.

Hull, C. L. (1951). *A Behavior System.* New Haven: Yale University Press.

Mace, C. A. (1935). Incentives: some experimental studies. *Rep. Med. Res. Coun. Industr. Hlth Res. Bd, Lond.* London: H.M.S.O.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 268. London: H.M.S.O.

O'Connor, N. & Claridge, G. S. (1955). The effect of goal-setting and encouragement on the performance of imbecile men. *Quart. J. Exp. Psychol.* **7**, 37–45.

Zeaman, D. (1949). Response latency as a function of the amount of reinforcement. *J. Exp. Psychol.* **39**, 466–83.

(*Manuscript received* 28 *December* 1956)

THE EFFECTS OF INCENTIVES ON THE PERFORMANCE OF DEFECTIVE IMBECILES

By D. WALTON

Rainhill Hospital, near Liverpool and Winwick Hospital, Warrington

and T. L. BEGG

Public Health Department, Liverpool

The present experiment was designed to inquire into the effects of different incentive conditions on the performance of male imbeciles. In particular the aim was to clarify further the effect of goal-striving and encouragement on performance in uncontaminated experimental groups. The experiment was suggested, first, because the previous results in the field had been equivocal, and secondly, since inquiry into the trainability of imbeciles had been encouraging. The latter results suggested how imperative it was to determine the best incentive conditions to be employed in any experimental workshop in which such work was being undertaken.

Forty-eight imbeciles were used. On the basis of their first performance on the Leg Persistence Test they were divided into four matched groups of 12 subjects each and asked to perform under different motivating conditions. There were four incentive conditions used—'Control', 'Goal', 'Competition' and 'Encouragement'. The experiment continued at the rate of one trial per day over twenty trials. As far as possible each of the four groups was tested on the same day for the equivalent trial number. The four groups were also found to be closely matched on the basis of intelligence.

There can be little doubt that incentives improve the performance of imbeciles on a dull routine task. In contrast, under conditions of indifference their performance deteriorates. No support was gained for Gordon's contention that imbeciles respond to goal conditions of motivation better than when encouraged. In the present study competition and encouragement appear to be the best conditions of motivation.

I. Introduction

With the publication of Gordon's work (1953) has come a revival of interest in the relatively unexplored field of how low-grade defectives respond to incentives. Although previously Aldrich & Doll (1931) and Gardner (1945) carried out experiments with imbeciles in which incentives were used, it is fair to say that this area of study has been relatively neglected. Such a pessimistic view of the abilities and trainability of imbeciles has been adopted by such authorities as Tredgold (1952) and Lewis (1929) that the study of incentives seemed hardly worth investigating, and somewhat irrelevant. Tredgold (1952), for example, states:

> As a result of training, a considerable proportion can be employed in such simple duties as sweeping and scrubbing floors; polishing brasses; weeding garden paths; collecting potatoes; helping in the laundry and so on. But they can only do these things under supervision. They are quite incapable of adapting themselves to anything out of the routine to which they have been accustomed, and it is doubtful whether the work they do from an economic standpoint is worth the supervision it entails. None of them can contribute at all appreciably toward the cost of their keep....

Again, Lewis (1929), writes:

> Imbeciles are incapable not only of earning an independent livelihood, but even of contributing materially to their own support. The best that can be expected of imbeciles will be the simplest of routine tasks under supervision. The brightest...can manage in a somewhat irregular fashion, such menial duties as sweeping, dusting; scrubbing floors; washing earthenware and unbreakable articles, and even rough laundry work; even here, however, they would need almost continuous supervision....In wood-work the best of the imbecile youths can manage rough-polishing, sand papering and hammering large nails, and in needle-work the best of the girls can manage coarse stitching and tacking; few will be able to thread their own needles...(and)...tying shoe laces often remains entirely beyond their powers.

It is only recently, however, that the social and psychological aspects of mental deficiency have come to be studied with the development of a systematic psychology of the defective as its principal aim. Amongst these studies have been investigations into the ability of imbeciles to learn skills and be employed doing simple contractual work in sheltered workshops (Tizard & Loos, 1954; Loos & Tizard, 1955; Clarke & Hermelin, 1955). The results from these studies throw doubt on traditional views about the ability and trainability of defective imbeciles. As Clarke & Hermelin (1955) remark: 'It seems that the limits to the trainability of imbeciles are very much higher than have been accepted traditionally either in theory or in practice.'

These early experimental results also suggest the importance of submitting to systematic inquiry, not only the abilities and trainability of defectives, but also the study of the best incentive conditions to be used for this type of patient in such experimental workshops. Contrary to the suggestions of Lewis (1948) and Kallman (1951) that imbeciles constitute a separate clinical entity and that the basic determinants of behaviour found in normals and even in high-grade feeble-minded patients do not apply to them, the work of Gordon (1953), Gordon, O'Connor and Tizard (1954, 1955), O'Connor & Claridge (1955), Tizard & Loos (1954), Loos & Tizard (1955) and Walton & Begg (1955) would suggest that there is no real difference between normals and imbeciles in their response to incentive conditions.

Gordon *et al.* (1954) found, for example, that imbeciles were capable of responding to incentives, and that when they were working toward a goal they did better than when simply encouraged. Both incentive groups did significantly better than the control group who were given the original instructions and nothing else. They concluded that in their manner of responding to incentive, imbeciles did not appear to differ from normals (Mace, 1935). Certain criticisms, however, could be made about Gordon's work. First, very small numbers of subjects were used, particularly in the later experiments. Secondly, the superiority of the 'Goal' group may have depended on Gordon's personal approach to the imbeciles striving towards a goal and their successes or failures rather than 'their own self-competition as an internalized drive'.

O'Connor & Claridge (1955) attempted to make a more clear-cut distinction between 'Encouragement' and 'Goal-striving'. For this purpose they used 'Goal with Encouragement' and 'Goal with Indifference'. Their results indicated that goal-striving by itself did not increase the performance level significantly unless it was accompanied by 'social approval of work' and encouragement.

Loos & Tizard (1955) asked the question: 'What conditions will enable them (i.e. defectives) to work steadily and happily throughout a normal working day without compulsion or punishment?' In the early stages of their investigation they found general confirmation of Gordon's findings in that when imbeciles are given goals to work for there is an improvement in their performance. In one important respect they did not find confirmation of his work. They found, for example, that after two weeks there was a gradual fall. Their conclusion was, 'that in a workshop our subjects' desire to beat their own best performance was not a strong enough motivational force to keep them working steadily throughout a normal working day—after the novelty of the change in conditions had apparently worn off'.

When the incentive conditions were changed in such a manner as to give a good reason why the defective should work fairly consistently there was improvement. In their third experiment the most favourable conditions were those in which the imbeciles worked as

members of small groups where their own production was related to high-grade patients, with whom they worked on a basis of equality. It suggested that the factor of self-esteem was of prime importance in both determining and maintaining an improvement in performance.

The present experiment was designed to clarify further the effects of different incentives on the performance of imbeciles and to determine more accurately the respective roles of goal-striving and encouragement in uncontaminated experimental groups.

II. METHOD

Forty-eight male imbeciles were used. Each subject was individually given the Leg Persistence Test (Gordon *et al.* 1954). First, the procedure was demonstrated to the subject. The subject was then seated with a chair opposite him and given the following instructions: 'When I tell you, I want you to keep your leg above this chair for as long as you can. Do not touch the chair with your leg but try to hold it up for as long as you can without touching the chair. Are you ready? Go!'

The subject's score was the number of seconds he held his leg up in the air.

On the basis of the scores from the first trial, the 48 subjects were divided into four matched groups of 12 subjects each, and asked to perform under different motivating conditions. Four incentive conditions were used: (1) *Control.* The subject was given the above instructions. No encouragement or further comment was made by the examiner. (2) *Goal.* The subject was able to follow his progress by watching a pointer record his persistence. Another pointer showed his level of achievement on the previous trial. As under control conditions, no encouragement or comment was made by the examiner beyond exactly the same instructions as given to the control group. (3) *Competition.* This group watched each other's progress, knowing their own results and the results of the other subjects in the group. As under 'Goal' conditions the 'pointer' was used. One 'pointer' showed the highest score on that day by a member of the group whilst the second recorded the persistence of the remaining subjects. No encouragement or comment was made by the examiner beyond the same instructions given to the control group and an explanation of what was demanded of them. (4) *Encouragement.* No 'pointers' or observable goals were available to the imbeciles in this group. They were given the same instructions as the control group, followed by continued personal contact and praise of the subject by the examiner during each of the experimental trials.

The experiment continued at the rate of one trial per day over a further twenty trials. Except for certain unavoidable delays each of the four groups was tested on the same day for the equivalent trial number. In all, there were 960 observations.

There was no significant difference between the four groups in their initial scores on the Leg Persistence Test. The means in seconds were respectively (Control, $\bar{x}=70{\cdot}0$; Goal, $\bar{x}=66{\cdot}2$; Competition, $\bar{x}=69{\cdot}6$; Encouragement, $\bar{x}=69{\cdot}1$). The mean Terman-Merrill I.Q.'s and S.D.'s for the Control, Goal, Competition and Encouragement groups were 47·54, S.D. 10·0; 47·90, S.D. 6·4; 45·0, S.D. 8·5; 46·58, S.D. 8·8 respectively.

III. RESULTS

There are several difficulties about the present data. If the experiment consisted of three treatments being applied to different groups of individuals, so that the differences in response obtained between the several individuals within each of the groups (including the fourth control group) all behaved like experimental errors (i.e. quite randomly and independently of the treatment applied), one could proceed to test the effects of the treatments by a simple analysis of variance. The trouble with the present data is that they do not seem to behave in this way.

In the first place, there are clearly as great differences in the variability from trial to trial within each of the groups as there are in their average levels. In fact, the differences between each of the four groups in their variances or standard deviations are very much greater proportionately than the differences in the four means. This must suggest that there is some sort of interaction between the factors being tested and the successive trials: i.e. that

the effect of the three factors on the response obtained cannot be assumed to be the same at each successive trial.

This can be seen from another point of view. The result obtained at each trial, at least in the control and competition groups, seems to be correlated with the result at the preceding trial. The serial correlations present in each of the four groups are given in Table 2. There is no evidence of second or higher order serial correlations except in the control group. All the figures are the coefficients squared, so the square root is needed if the actual correlations are wanted. There is, therefore, evidence of trend or cyclical fluctuation in the results obtained, especially in the control and competition groups. This can be seen quite clearly from Fig. 1, especially at the beginning of the experiment. Of course, one cannot treat 'competition' as if it were a *constant* factor influencing the results obtained, since it must to some extent be variable from trial to trial.

Some may argue, therefore, that it is not appropriate to use the analysis of variance, since the conditions required to make valid tests of significance do not seem to be met by the present data. This is not in the least serious since the obvious way to handle the data is to have a good look at how it behaves, especially by plotting the results graphically, and to see how one can interpret the results best, using common sense and the specialized knowledge one possesses. Significance tests are not, and never have been, a substitute for inspecting the data thoroughly oneself to see what conclusions can be drawn. Often they are either redundant, because the differences are patently significant merely from looking at the data, or they cannot decide doubtful cases because there is some doubt whether the data really meet the requirements and assumptions underlying the tests.

In the present case, the differences between the means of the four groups are obviously quite significant. Indeed, if one carries out an analysis of variance irrespective of the above objections, the F ratio works out at 83·89 which (assuming the validity of the tests) is highly significant. The differences between trials do not seem to be significant. The only thing in doubt is whether the difference between the competition and encouragement group is significant. A t-test ('t' = 1·8698, $0{\cdot}10 > P > 0{\cdot}05$) suggests it is not, but the t-test is open to the same objections as the analysis of variance.

From an inspection of the data several trends are suggested. First, the experiment lasted one month and bears comparison with the work of Loos & Tizard (1955). Both studies reflect the superiority of incentive conditions, but both studies also agree in that when imbeciles are set relatively routine tasks, even under incentive conditions, their performance does deteriorate, though it never reaches such a low level as under control conditions. The results from the control group are of particular significance in relation to what was said at the beginning of the paper. Under conditions of indifference their performance declines. In the present study this started immediately. The work of Tizard (1953) is also of relevance. He found that supervision with indifference produced poor results compared with supervision involving encouragement. Early in the present experiment subjects in the control group complained a great deal about the boring nature of the task. This also spread to the other groups at a later point in the experiment. It is important, however, to make a special point that this group of imbeciles, though of minimal intellectual endowment, appeared to become very disgruntled at the prospect of an uninteresting task before them. This appears to cut right across current theory. Of particular interest was the distractibility of the subjects in the competition group. Some wandered away and started playing with a bat and ball in a far corner of the room.

No support was found for Gordon's suggestion that imbeciles respond to goal conditions of motivation better than when encouraged. A previously stated criticism of his work—that he may have inadvertently encouraged the subjects and thereby have contaminated his goal-group results—appears justified by the relative superiority of the 'Encouragement' group results in the present experiment. The implication of the study carried out by Loos & Tizard (1955), that the factor of self-esteem was of great importance in determining and maintaining an improvement in performance, was however confirmed. The upshot of the results is that when imbeciles are given an uninteresting task to perform their level of achievement can be improved significantly by the use of incentive conditions. In the

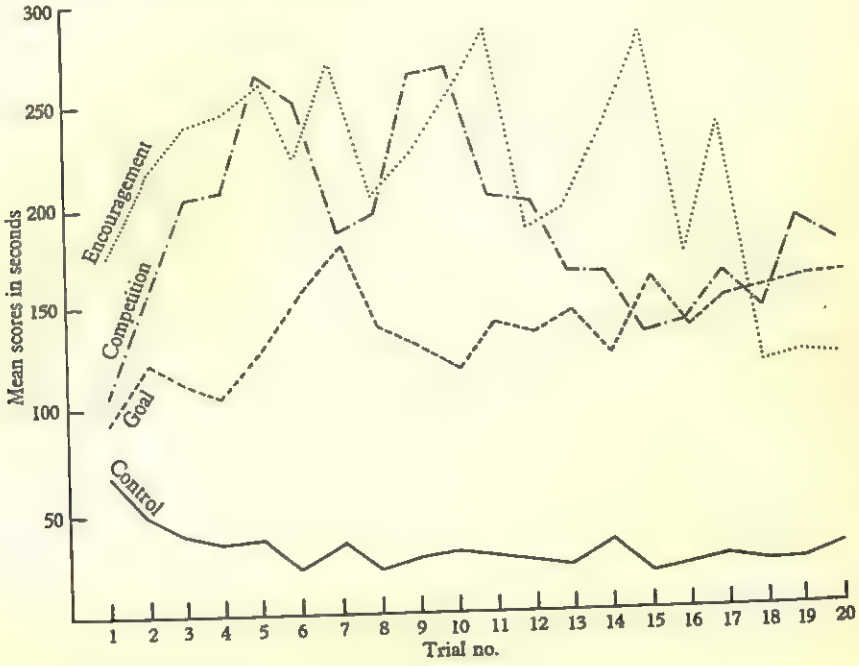


Fig. 1

Table 1. *Group mean scores (in seconds) under experimental and control conditions*

Trial	Control	Goal	Competition	Encouragement
1	68·8	94·8	105·0	176·0
2	50·2	121·0	154·0	219·0
3	44·7	114·0	204·0	239·0
4	35·7	106·0	208·5	244·8
5	36·2	128·7	266·0	261·3
6	23·5	159·5	250·0	220·2
7	36·0	182·1	185·0	268·1
8	23·1	142·0	190·0	202·0
9	30·0	132·5	263·0	221·0
10	32·1	118·0	265·0	250·0
11	30·7	143·0	201·5	283·0
12	27·0	137·9	199·5	177·0
13	26·7	146·3	168·0	192·0
14	35·1	124·8	168·0	230·0
15	18·5	160·8	131·0	281·0
16	21·0	137·5	135·0	175·0
17	26·8	150·0	166·5	238·0
18	24·0	152·0	146·7	117·5
19	24·0	156·8	189·0	123·5
20	29·5	160·7	176·0	121·8

present experiment, 'goal', 'competition' and 'encouragement' conditions of motivation produced significantly better results than under control conditions. The level of achievement under control conditions undergoes a progressive deterioration. Competition and encouragement seem to provide superior motivation conditions than does goal-striving.

Table 2. *Means, standard deviations, variance and serial correlations for experimental and control conditions*

	Control	Goal	Competition	Encouragement
Mean	32·18	138·42	188·58	212·01
S.D.	11·35	20·84	44·56	49·56
Variance	128·89	434·42	1984·74	2455·72
1st order serial Correlation	0·474213	0·235826	0·383989	0·1160805
2nd order	0·465343	0·053813	0·048011	0·068185
3rd order	0·227132	0·022548	0·008888	0·020756
4th order	0·145306	—	—	—

Table 3. *Analysis of variance table*

Variance	D.F.	Sum of squares	Mean squares	*F* ratio	Level of significance
Between groups	3	382,845·18	127,615·06	83·89	0·001
Between Trials	19	28,900·62	1,521·085	1·218	Not significant
Interactions	57	71,174·87	1,248·682		
Trials and interactions	76	100,075·49	1,316·783	—	—
Total	79	482,920·67	61,129·20	—	—

IV. Implications

It has been suggested elsewhere (Lewis, 1948; Kallman, 1951) that imbeciles constitute a separate clinical entity and that the basic determinants of behaviour found in normals do not apply to them. From a comparison of Mace's work (1935), using normal subjects, and the results of the present study plus the work of Gordon (1953), Gordon *et al.* (1954, 1955), O'Connor & Claridge (1955), Tizard & Loos (1954) and Loos & Tizard (1955), all using imbeciles, there appears little or no support for the latter contention. Self-esteem seems to be one factor common to normal and imbecile subjects in determining level of behaviour.

Not only do some of the basic determinants of behaviour appear to suggest themselves as the same in imbeciles as in normals, but recent work is tending to show also that the potential level of achievement in imbeciles is considerably higher than suggested by such authorities as Lewis (1929) and Tredgold (1952). The work of Tizard & Loos (1954) and Clarke & Hermelin (1955) is of particular relevance. The latter authors have shown, for example, that imbeciles can work for long periods and contribute significantly to their keep. Six imbeciles were employed for two and a half years in an experimental workshop as cardboard-box folders. Their average weekly output was 30,000–40,000 for a 35-hour week. Others were able to earn £2 to £3 per week cutting insulated wire to exact lengths using a simple guillotine, others were taught 'soldering of four different coloured wires to the correct terminals of an 8-pin television plug. This requires not only soldering ability but also the distinction of colours and spatial relationship.' Both studies also show that face-value impressions or a subject's initial scores are likely to give a very poor idea of his ability to do tasks after practice. As Clarke & Hermelin (1955) say: 'It seems that to take the imbecile at "face-value" is to ignore his potentialities. These would enable many, after

training, to perform certain useful tasks and to contribute in many cases substantially to their own support.'

The third fact, which has been noted not only in the present study, but elsewhere, is that when imbeciles are set routine, uninspiring tasks to do without incentives, or are left to their own devices without previous training, their performance deteriorates. Clarke tells of a weekend when, in the absence of the supervisor, all six imbeciles from the workshop became bored and gained entrance into the closed workshop and worked all morning without supervision. If one imagines, therefore, a group of hospitalized imbeciles allowed and expected only to perform tasks similar to those suggested by Tredgold and Lewis, then it is not surprising to find that face-value impressions appear correct. As mentioned above, however, controlled experiments have shown that the initial level of performance gave no idea of the ultimate level to which the defective may rise.

If, therefore, imbeciles have been seen to respond favourably to incentives, that they are capable of acquiring 'skills' considerably in excess of those previously designated to them, it seems indefensible to allow them to be simply occupied—'... a more positive approach is needed here quite as much as in other spheres of remedial treatment. Just as, in many countries, rehabilitation centres and sheltered workshops have been established for people with severe physical handicap whose condition was formerly considered hopeless, so it should be possible to aid many of the mentally handicapped who can never be socially independent but who with similar help could contribute substantially to their own support (Clarke & Hermelin, 1955).'

References

Aldrich, C. G. & Doll, E. A. (1931). Problem solving among idiots. *J. Comp. Psychol.* **12**, 137–169.

Clarke, A. D. B. & Hermelin, B. P. (1955). Adult imbeciles. Their abilities and trainability. *Lancet*, Vol. **2**, 337–9.

Gardner, L. P. (1945). The learning of low-grade aments. *Amer. J. Ment. Defic.* **50**, 59–80.

Gordon, S. (1953). Some effects of incentives on the behaviour of imbeciles. Unpublished Ph.D. thesis, University of London.

Gordon, S., O'Connor, N. & Tizard, J. (1954). Some effects of incentives on the performance of imbeciles. *Brit. J. Psychol.* **45**, 277–87.

Gordon, S., O'Connor, N. & Tizard, J. (1955). Some effects of incentives on the performance of imbeciles on a repetitive task. *Amer. J. Ment. Defic.* **60**, 371–7.

Kallman, F. J. (1951). Recent progress in relation to the genetic aspects of mental deficiency. *Amer. J. Ment. Defic.* **56**, 375–81.

Lewis, E. O. (1929). Report of an investigation into the incidence of mental defect in six areas, 1925–1927. *Report of the Mental Deficiency Committee*, Part iv. London: H.M.S.O.

Lewis, E. O. (1948). Mental deficiency and social medicine. *J. Ment. Sci.* **94**, 258–65.

Loos, F. M. & Tizard, J. (1955). The employment of adult imbeciles in a hospital workshop. *Amer. J. Ment. Defic.* **59**, 395–403.

Mace, C. A. (1935). *Incentives: Some Experimental Studies.* London: H.M.S.O.

O'Connor, N. & Claridge, G. S. (1955). The effect of goal-setting and encouragement on the performance of imbecile men. *Quart. J. Exp. Psychol.* **7**, 37–45.

Tizard, J. (1953). The effects of different types of supervision on the behaviour of mental defectives in a sheltered workshop. *Amer. J. Ment. Defic.* **58**, 143–61.

Tizard, H. & Loos, F. M. (1954). The learning of a spatial relations test by adult imbeciles. *Amer. J. Ment. Defic.* **59**, 85–90.

Tredgold, A. F. (1952). *A Textbook of Mental Deficiency.* London: Baillière, Tindall and Cox.

Walton, D. & Begg, T. L. (1955). Adult imbeciles. *Lancet*, Vol. 2, 616–17.

(*Manuscript received* 28 *December* 1956)

A 'CRESPI EFFECT' IN THE ANALYSIS OF EMOTIONALITY AS A DRIVE IN RATS

By P. L. BROADHURST

University of London, Institute of Psychiatry (Maudsley Hospital)

In a previous paper (Broadhurst, 1957), it was shown that, when air deprivation is used as a drive, emotional rats swim significantly faster than non-emotional ones. This effect, which had been predicted by analogy with Eysenck's formulation of neuroticism as an autonomic drive (1955, 1956), was interpreted as reflecting a stronger drive in the emotionally reactive animals. It was not however possible, in the situation used, to decide between two of the possible explanations of the effect observed—a generally higher level of drive, or a greater tendency to drive-arousal on the part of the emotional animals. A simpler experimental situation in which the level of the imposed drive is varied, and which has given rise to the phenomenon termed the 'Crespi effect' was therefore used. Crespi (1942, 1944) showed that when rats who had been trained to run to a small food incentive (low drive) were suddenly rewarded with a larger one, they increased their speed of running to a point beyond that reached by subjects who had been trained from the start under the high drive (large incentive) condition. This was called the 'elation' effect; the converse 'overshoot' in the opposite direction in response to a sudden reduction in drive was termed the 'depression' effect. Crespi's interpretation in terms of emotional drives makes this situation a fitting one for the evaluation of differences due to emotionality. Zeaman (1949) used latency of response in a similar situation, and found both of the Crespi contrast effects when the amount of the food reward was suddenly changed. The depression effect did not reach significance, however. Deese & Carpenter (1951), in a similar study of changes in the latency of response to alterations in drive strength, failed to provide controls by which the Crespi contrast effects could be evaluated. It is, however, clear from their data that the depression effect is absent. The wide generality of the elation effect is shown by the recent finding by O'Connor & Claridge (1958) that imbeciles working at a repetitive task show a significant elation effect in response to a sudden increase in incentive in the form of encouragement and goal setting. These workers also failed to find the depression effect.

The elation effect only will therefore be used in the present experiment, the subjects for which were 48 albino rats, selected from the second and third generations of a selective breeding study of emotionality (Broadhurst, 1958*b*). They had been reared under standard conditions, before their emotional reactivity was observed in Hall's open-field test (1934), under the standardized conditions described elsewhere (Broadhurst, 1958*a*), when they were 104·9 days of age ($\pm$S.E. 0·35). By virtue of their defecation scores in this test they were assigned to an emotional group (mean score $4{\cdot}23 \pm 0{\cdot}55$ fecal boluses per day) or to a non-emotional group (mean $0{\cdot}61 \pm 0{\cdot}54$). This difference in mean scores is significant beyond the 0·1% level by *t*-test. Subsequent to this testing, but immediately prior to the present experiment, they had served as subjects for a learning task in a Y-shaped underwater discrimination unit, for which four different air deprivation drive levels had been established and used for 15 consecutive days, by submerging the subjects for 0, 2, 4, or

8 sec. under water before release for each learning trial (Broadhurst, 1957). There were thus 12 subjects, six emotionals and six non-emotionals, habituated to swimming to escape from underwater at each of four different levels of drive.

In the present experiment a 10 ft. long straight tank, 6 in. wide, was used. The water-level was 9 in., its temperature was maintained at 20° C. Air temperature averaged 20·8 (±0·40)° C. A submersible starting cage with a guillotine door which could be opened from above the surface occupied the first 1 ft. of the length of the tank, a wire-netting ramp rising at an easy angle from below the surface occupied the last 1 ft. The intervening 8 ft. were roofed in with stainless steel wire gauze $\frac{1}{2}$ in. below the surface. The subjects were thus forced to swim 8 ft. underwater in order to escape. The time from the release of the door of the starting cage until the subject's snout broke surface at the other end of the tank was recorded to the nearest $\frac{1}{10}$ sec. by stop-watch. These times directly represent speed of swimming since the distance swum is constant. With the same four air deprivation drive levels as previously used, one day's preliminary training was given in this long tank by gradually lengthening the underwater 'ceiling' on successive trials, and on the next day the experiment began with 10 trials of swimming the whole 8 ft. The inter-trial interval was held constant by running those subjects subjected to short delay times in larger groups than those subject to longer ones. On the second day of the experiment the motivation was increased by subjecting all animals to the same delay as that imposed on the two 8 sec. groups, whose motivation was unchanged, and who consequently provided control groups. All animals were given a further ten trials on the next day under the same conditions, and half of each group, randomly selected, another ten on the fourth day. These latter results are shown in Fig. 1, but not included in the subsequent analyses.

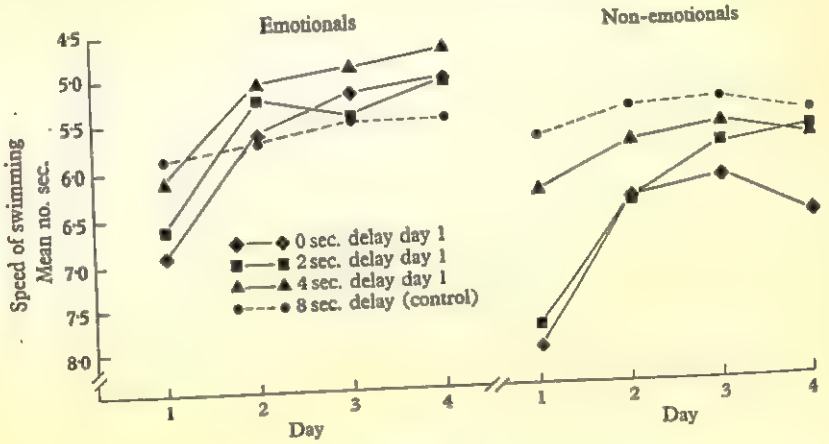


Fig. 1

It is clear from Fig. 1 that the Crespi elation effect is present, but among the emotionally reactive rats only. There is some suggestion that the amount of performance 'overshoot' is a function of the magnitude of the change in drive level. The effect appears to persist, or even increase a little, especially among the emotional rats. In order to investigate inter-group differences in level of drive, a 2×4 analysis of variance of the average times each animal returned on the day before the increase in drive (day 1) was performed, after it had been shown that the variance of these data was satisfactorily homogeneous (Pearson & Hartley, 1956). This analysis showed that the F ratios for the two main effects of emotionality and motivation level were significant, with no significant interaction between them. The effect of the four different levels of drive is highly significant, as might be expected ($P < 0{\cdot}001$); the emotional rats swam significantly faster than the non-emotionals ($P < 0{\cdot}05$). While this is a relatively lenient criterion, the whole analysis so precisely confirms the findings previously reported (Broadhurst, 1957) that it may be

concluded that emotional rats show a significantly higher drive level than non-emotional ones, as evidenced by their respective speeds of swimming.

In order to investigate the arousal of drive in response to the sudden increase in motivation among the groups previously delayed for 0, 2, or 4 sec., a score reflecting the increase in speed (decrease in time) was computed by subtracting each animal's average time on the second day from that for the first. Each score so obtained was corrected by deducting the average decrease for the six rats in the appropriate control group, similarly calculated. This correction was necessary because the control groups (8 sec. delay throughout) were manifestly also improving their speed (see Fig. 1). A second analysis of variance, like the first, was then performed on these data and revealed no significant effect whatsoever. There was a tendency for the emotionals to increase their swimming speed more than the non-emotionals (mean decrease in swimming time 1·09 and 0·89 sec. respectively), just as there was a consistent trend for a larger increase the lower the original motivation, but neither reached significance. We may therefore further conclude that, within the limits of this experiment, emotionally reactive rats do not display a significantly greater arousal of drive, as measured by an increase in their speed of swimming, than do non-reactive ones, when both groups are subjected to an increase of motivation.

References

Broadhurst, P. L. (1957). Emotionality and the Yerkes-Dodson Law. *J. Exp. Psychol.* **54**, 345–52.

Broadhurst, P. L. (1958*a*). Determinants of emotionality in the rat. II. Antecedent factors. *Brit. J. Psychol.* **49**, 12–20.

Broadhurst, P. L. (1958*b*). Studies in psychogenetics: the quantitative inheritance of emotionality in the rat investigated by selective and cross-breeding. *Bull. Brit. Psychol. Soc.* **34**, 3A. (Abstract.)

Crespi, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.* **55**, 467–517.

Crespi, L. P. (1944). Amount of reinforcement and level of performance. *Psychol. Rev.* **51**, 341–57.

Deese, J. & Carpenter, J. A. (1951). Drive level and reinforcement. *J. Exp. Psychol.* **42**, 236–8.

Eysenck, H. J. (1955). A dynamic theory of hysteria and dysthymia. *J. Ment. Sci.* **101**, 28–51.

Eysenck, H. J. (1956). Reminiscence, drive and personality theory. *J. Abnorm. Soc. Psychol.* **53**, 328–33.

Hall, C. S. (1934). Emotional behavior in the rat. I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. Psychol.* **18**, 385–403.

O'Connor, N. & Claridge, G. S. (1958). A 'Crespi effect' in male imbeciles. *Brit. J. Psychol.* **49**, 42–48.

Pearson, E. S. & Hartley, H. O. (1956). *Biometrika Tables for Statisticians*, vol. I. Cambridge University Press.

Zeaman, D. (1949). Response latency as a function of the amount of reinforcement. *J. Exp. Psychol.* **39**, 466–83.

(*Manuscript received* 24 *January* 1957)

A PATTERN RECOGNITION METHOD OF LEARNING MORSE CODE

By M. D. ALLAN

Senior Psychologist at the Branch of Scientific Advisor to Air Ministry

This study is concerned with a comparison between the traditional 'analytic' method of learning the morse code, and a 'Gestalt' approach called the Pattern Recognition method. The Pattern Recognition technique consists of transmitting characters right from the start at a speed corresponding to 20 words per minute. There is no preliminary memorizing of the alphabet. All the pupils have to do is to listen to the whole character (i.e. a letter or a number) transmitted at 20 words per minute and, when the instructor has written its meaning on the blackboard, to copy it down. This sort of drill is continued until all the 36 characters (26 letters and 10 numerals) are instantly recognized by their distinctive sound patterns. The pilot experiment showed that pupils trained by the Pattern Recognition method had a significantly better knowledge of the alphabet, after the same amount of training, than pupils trained analytically. As an applied technique in a short intensive course, the Pattern Recognition method produced significantly fewer unsatisfactory pupils during the training period. In a long course, where pupils practised at whatever speed they were capable, pupils trained by the Pattern Recognition method arrived at high speeds earlier than those trained analytically.

I. Introduction

Morse reception is a speed skill involving the ability to recognize and name auditory patterns. The patterns are composed of different combinations of short and long sounds, each symbolizing a letter or a numeral. In teaching such a skill, a choice must be made between two main varieties of learning theory—the characters can be studied and learned analytically, or they can be treated as 'Gestalten' or indivisible patterns of sound. The oldest method of teaching the morse code, and one still in common use, is an analytic method. The student is first given a card in which all the letters of the alphabet and the ten numerals are represented as visual dots and dashes. He is expected to study this analytically and memorize the parts. The characters are then tapped out slowly by the instructor so that the student can distinguish clearly the 'dits' and 'dahs' which he hears, and match them up with the dots and dashes which he has already memorized. The speed of transmission is increased gradually from day to day until the required speed of reception is attained. This change of speed involves changes in the duration of the elements because everything must be scaled to conform with the international agreement that: (*a*) a dot is one unit, (*b*) a dash is equal to three units, (*c*) the space between component parts of a character equals one unit, and (*d*) the space between any two characters equals three units. By this method, the proportional relationship between the elements remains constant throughout the learning period, although the duration of the separate long and short sounds changes as the speed of transmission changes. It is difficult to say to what extent this alters the sound of individual characters. Certainly, the auditory impression of a character transmitted in this way at 4 words per minute must be very different from the auditory impression of the same character transmitted at operational speed—say 20 words per minute.

The fact that the morse code is constructed entirely of different combinations of short and long sounds makes the analytic method very attractive—it straightaway suggests

the memorization of components. Yet a morse character is a distinctive pattern of sound —it is a whole sound, and it must be experienced immediately as such before operational speed is possible. The skilled operator has no time to count up the separate elements, so at some point during the learning process he must stop listening to morse characters analytically and instead experience the sound patterns as wholes. This consideration suggests that it might be profitable to forget that morse characters can be so easily analysed, and ask instead if it is possible for beginners to be trained from the start to recognize and name 36 different patterns of sound (i.e. 26 letters and 10 numerals) by experiencing the sound only. This is the 'Gestalt' approach, and it demands that individual morse characters should be regarded as indivisible patterns of sound and transmitted at a constant high speed right from the start. Slow speeds are obtained by increasing the spaces between characters.

II. Outline of previous research

The idea of learning individual characters at high speed was first suggested in 1918 in a bulletin issued by The Federal Board for Vocational Education (Washington), but appears to have passed unnoticed, probably because there was no supporting evidence. In 1919, Lipmann discovered that an expert operator found it difficult to prolong dots and dashes at low speed to conform with the fixed proportional relationship between elements cited above, and in practice he tended to decrease his speed by increasing the gap between characters. This led Lipmann to suggest (1928) that individual characters should be transmitted at a constant rate of 20 words per minute and slowness for beginners obtained by increasing the spaces between characters. So far, no experimental evidence existed to support these recommendations, but Biegel (1932) and Koch (1936), recognizing that a morse character at high speed has a distinctive rhythmic character, investigated the matter from the point of view of Gestalt theory. These investigators produced some impressive, but inconclusive, experimental evidence in support of the view that a morse character should be treated as an indivisible pattern of sound. The first attempt to compare the relative merits of the 'Analytic' and the 'Gestalt' approach to morse training was made by Taylor (1943*b*) in a carefully controlled experiment involving 13 subjects covering a training period of 40 hours. His findings showed no significant difference between the two methods in speed of learning.

The generally accepted training in the Royal Air Force has for years been the analytic form and, since no complaint could be made about the standard of skill eventually achieved by operators, probably no change would have been contemplated. The matter, however, came under consideration in the Royal Air Force because of complaints that pupils experience, during the early stages of training, a degree of strain out of proportion to the difficulty of the task. There were many complaints of headaches during training, and the Service records showed a fairly high casualty rate due to apparent lack of aptitude. It occurred to the writer that the strain may be caused by pupils analysing when they ought to be listening. The analytic method demands a summation of elements, and the effort to do that at increasing speeds must impose considerable strain. Furthermore, the fact that the duration of dots and dashes changes with the rate of sending means that the sound of individual characters must change as speed increases. Consequently, during the learning period, continual adaptation must be made to changes in auditory experience,

although such adjustment is not a feature of the learned skill. These considerations suggested that it might be profitable to re-examine the recommendations of Lipmann, Biegel and Koch that speed of transmission of individual characters should be held constant, and that variations in speed should be obtained by varying the interval between characters. Thus, at low speeds, the individual letters or numerals would be transmitted at a constant high speed but the interval between each would be lengthened to produce the over-all low speed. Permission was obtained for the matter to be investigated on Royal Air Force personnel in a live training situation, on the understanding that the normal training schedule would not be interrupted.

III. PLAN OF INVESTIGATION

The plan of investigation consisted first of a pilot study, the purpose of which was to find out if pupils were able to learn morse characters by the method of direct presentation, and if so, to obtain some measure of how effective the learning was, compared with a more orthodox method.

(1) *The Pilot Experiment*

A class of 64 airmen about to start training as wireless operators were taught the morse alphabet by the Pattern Recognition method: that is, individual morse characters were sent out by automatic transmitter at a speed corresponding to 20 words per minute with an interval of 7 seconds between each character. Both the instructor and pupils listened and, as each letter came over, the instructor wrote it on the blackboard, and the pupils copied it down. In writing up the letters, the instructor used 5-letter groups so that the pupils should get used to the form of the 5-letter code. Thus, right from the start, pupils had to listen to whole characters, and they were not permitted to memorize analytically either from a card or from studying visual symbols. The idea was to develop recognition and to establish, by mechanical means, an association between the sound pattern and its meaning which would lead eventually to automatic identification of the sound.

Table 1. *Test of knowledge of letters in 6 words per minute test*

	Pass	Fail	χ^2	P
Pattern Recognition method, $N = 66$	51 (77%)	15 (23%)	16·1	0·001
Analytic method, $N = 71$	31 (44%)	40 (56%)		

To test progress a comparison was made between this group and a control group of 71 men trained by analytic techniques on the same station after the same amount of training. The test was on mixed letter groups at 6 words per minute, which is not really a test of speed but a test of knowledge of the alphabet.

The highly significant result (see Table 1) not only showed that the pupils were learning the code by direct presentation, but that their knowledge of the alphabet was better, after the same amount of training, than that of analytically trained pupils.

(2) *Applied experiments*

The next step was to incorporate this Gestalt approach into a complete training scheme and try it out experimentally against matched control groups trained by the analytic method. This was done at two different R.A.F. units, one on a short intensive course, and the other on a long course where morse instruction is spread out over 18 months.

(*a*) *Short intensive course.*

The object of the course is to train operators to receive and send morse in plain language, letter code, and letter/figure code at 20 words per minute.

A sample of 41 men were trained throughout by the Pattern Recognition technique and their progress compared with a control group of 39 men trained in the ordinary analytic way. A direct comparison in terms of pass/fail results in progress tests was inconclusive because, in the interests of training economy, some reshuffling of personnel takes place quite early in training. Pupils who are not making satisfactory progress are either suspended or reclassed for further training. A pass/fail comparison was made up to

the speed of 11 words per minute, and this is shown in Table 2. So far as they go, these results indicate that the ability to cope at low speeds is about the same whichever method is used, but that the Pattern Recognition method makes higher speeds easier. A more decisive result is obtainable when the groups are compared in regard to losses sustained through inability to cope with morse. The losses may not be absolute losses, but they do indicate unsatisfactory progress for, in addition to 'Suspensions' they include 'Reclassing' for reasons involving unsatisfactory progress in morse. A comparison in these terms was made covering the whole training period up to the level of operational skill, and it was found that the Pattern Recognition group had significantly fewer losses ($\chi^2 = 6{\cdot}15$, $P < 0{\cdot}02$).

Table 2. *Comparison of training progress in standard tests up to 11 words per minute*

Speed	Pattern Recognition method, $N = 41$		Analytic method, $N = 39$		P
	Pass	Fail	Pass	Fail	
8 w.p.m.	30 (73%)	11	30 (77%)	9	0·90
9 w.p.m.	29 (71%)	12	23 (59%)	16	0·30
10 w.p.m.	32 (78%)	9	24 (62%)	15	0·10
11 w.p.m.	33 (80%)	8	21 (54%)	18	0·01

Table 3. *Comparison of training progress in terms of practising speeds*

	18 w.p.m. and over	15 w.p.m.	12 w.p.m.
Experimental group (Pattern Recognition), $N = 38$	14 (37%)	15 (39%)	9 (24%)
Control group (Analytic), $N = 44$	2 (5%)	23 (52%)	19 (43%)

$\chi^2 = 13{\cdot}86$; $P = 0{\cdot}001$.

(*b*) *Long course.*

Again it was not possible to make comparison of final pass/fail training results because, after such a long training period, very few pupils fail. However, it is the procedure at this particular school to allow pupils to practise at whatever speed they are capable. Here was a basis for comparison, provided that the comparison was made not too near the final examination where differences tend to be levelled out. In Table 3, the two groups are compared at a stage where they are normally expected to be receiving at 12–15 words per minute. The difference has high statistical significance (for 2 degrees of freedom, $\chi^2 = 13{\cdot}86$, $P = 0{\cdot}001$) and is further evidence that the Pattern Recognition method makes the learning of morse easier.

(3) *Further observations*

In addition to operative skill, pupils must also qualify in technical knowledge. Some indirect evidence that the Pattern Recognition method involved less strain was obtained when two of the groups were compared for technical knowledge. There was no statistical significance between the groups in terms of their intelligence ratings, yet in the final examination the group which learned morse by the Pattern Recognition method was superior to the control group at the 5% level of significance for 3 degrees of freedom. It is difficult to explain this since both groups had the same sort of instruction and the same standard of examination. It had been hoped that the new method would make learning morse less fatiguing on the operative side, so a possible explanation is that the pupils in the experimental group were less fatigued, and therefore more able to absorb theoretical knowledge.

IV. The effect on sending

Sending has never presented the same difficulties as receiving, so no special study was made of sending problems other than to see if the Pattern Recognition method of learning to receive would affect adversely learning to send. Normally pupils practise sending at very slow speeds at the same time as they practise receiving at slow speeds. The 'Gestalt' approach, however, rests upon preserving the distinctive wholeness of the characters, so formal practice in sending was postponed until pupils were thoroughly familiar with

receiving characters as wholes. During that time, however, pupils were shown the correct stance and method of holding the key, and were encouraged to fill in odd moments with flexibility exercises and even to try to reproduce with the key, during the 7 seconds interval, the patterns transmitted on the receiving tapes. Since no comparative studies were made on sending, nothing can be reported other than the fact that the method had no noticeable adverse effect on learning to send.

V. Discussion

The results of these experimental field studies point clearly to the usefulness of the Gestalt approach in learning what is, in fact, a recognition skill. The full benefits are more evident in the later stages of training than at the beginning. Perhaps this explains the non-significant results obtained by Taylor when he compared the two methods. His pupils were trained for 40 hours which, in an ordinary training scheme, is about the length of time required to bring the average beginner to a speed of 8 words per minute. There are limits to the extent to which the learning of an automatic skill can be speeded up by any method, and it seems, from the present study, that Taylor may have stopped his experiment too soon. The comparisons which are recorded in this paper began after about 40 hours of practice, and there was found to be little difference between the groups in the early comparisons, but the differences became significant as the skill approached operational level.

As well as making morse easier at higher speeds the Pattern Recognition method puts less strain on the learner. This could only be assessed indirectly as evidenced by the higher performance of the experimental group in technical knowledge (see § III (3) above) but it was very apparent to instructors who described their pupils as being 'less worried than usual'.

The method of instruction under discussion has been described as a 'Gestalt' approach, but it would be wrong to read into this more than the fact that the material is presented to the pupil in the form of whole patterns. Whether or not they react strictly in accordance with Gestalt theory is not at all clear. In the experiments, it was observed that they quickly became aware of the distinctive pattern of the sounds they were listening to; this was evidenced by the sharp reaction of the whole group when a stray character was introduced during one of the early lessons. But when it came to distinguishing one pattern from others and identifying it, there was evidence that many pupils were attending to some detail in relation to the whole pattern. This is a very different process from the memorizing of details which characterized the analytic method. There the details were wholes in their own right and the learning process was an additive one. In the Pattern Recognition method attention to detail, when it was found to occur, was always relative to the whole structure, and was more in the nature of selecting cues which would help to fix the identity of the whole. This brings the study into the realm of perception, for even a synthetic skill such as morse undoubtedly involves auditory perception. Intrinsically there is little to distinguish it from any of the recognition skills of everyday life which are acquired without formal training. These everyday skills, such as recognizing tunes or birds' songs, may involve much knowledge of detail, but the recognition process, where it is an involuntary one, always starts from experiencing the whole.

Investigators in the field of industrial operations have recently been drawing attention

to the perceptual element involved in speed skills and indicating that it is the more difficult element to master (Seymour, 1956). The present study is primarily directed towards solving the perceptual difficulties of morse, and it would appear that the repeated association of whole patterns of sound with their meaning allows free perceptual organization to take place on a subliminal level, rather characteristic of the way the recognition skills of everyday life are learned.

References

Anon. (1918). Emergency war training for radio mechanics and radio operators. *Bulletin Federal Board for Vocational Education*, no. 16. Washington: Government Printing Office.

Biegel, R. A. (1932). Das Anlernen der Höraufnahme durch Funkentelegraphisten. *Psychotech. Z.* **7**, 147–51. (Reviewed by Taylor, 1943, *Psychol. Bull.* **40**, 468.)

Koch, L. (1936). Arbeitspsychologische Untersuchung der Tägigkeit bei der Aufnahme von Morzezeichen, zugleich ein neues Anlernverfahren für Funker. *Z. angew. Psychol.* **50**, 1–70. (Reviewed by Taylor, 1943, *Psychol. Bull.* **40**, 469–71.)

Lipmann, O. (1919). *Z. angew. Psychol.* **15**, 301–40. (Taylor, 1943, *Psychol. Bull.* **40**, 468.)

Lipmann, O. (1928). Eignungsprüfungen für Funker. In Abderholden, E., *Methoden der experimentellen Psychologie*. Berlin: Urban und Schwerzenburg. (Taylor, 1943, *Psychol. Bull.* **40**, 468.)

Seymour, W. D. (1956). Experiments in the Acquisition of Industrial Skills. *Occup. Psychol.* **30**, 103.

Taylor, D. W. (1943*a*). Learning Telegraphic Code. *Psychol. Bull.* **40**, 461–87.

Taylor, D. W. (1943*b*). The learning of Radiotelegraphic Code. *Amer. J. Psychol.* **56**., 319–53.

(*Manuscript received* 13 *February* 1957)

THE DEVELOPMENT OF INDEPENDENT MUSCULAR CONTROL OF SEPARATE EYELIDS IN TWO RACIAL GROUPS, EUROPEAN AND POLYNESIAN

By P. R. N. SUTTON

University of Melbourne

The ability to control separate eyelids independently ('the voluntary winking ability') is studied in 527 subjects of European stock and in 262 Polynesians, with the following results: (1) In both groups, there is a significant correlation between voluntary winking ability and age. (2) Winking ability of the males is not significantly greater than that of the females. (3) In children under 10 years of age, the incidence of non-winking is significantly higher than in the remaining subjects. (4) In the European group, but not in the Polynesian, the incidence of those who could wink both eyes readily and with approximately equal facility is significantly lower in those who are 10 years and under, than in subjects over 10. (5) In those subjects of European stock who could wink one eye only or could wink one eye more easily than the other, a preponderance of the use of the left eye for winking is found. This is mainly due to the preference for winking the left eye which is seen in those who are under 10 years. However, in the Polynesians there is no significant difference between the frequency of winking each eye.

I. Introduction

Voluntary unilateral winking was considered by Duke-Elder (1952) and by Danielson (1932) to be an act which is usually established early in life and which it is difficult to change. The ability to wink the eyes independently has been observed in 789 subjects, 527 of them being of European stock and 262 of Polynesian, the age range in both groups being from 3 years to over 25 years, and some associations have been observed of which no record has been found in the literature. The results obtained in the two racial groups in regard to the ability to wink one eye voluntarily at a time are compared in this paper; and, for each group, the influence of the age and sex of the subjects is considered.

The influence of sex and of age on winking ability. W. Rider (1898) found that in 'older persons...there is still a small residuum of non-winkers, eleven out of 1343 men and seventy-three out of 626 women'. He felt that this 'disproportion between the sexes in the matter of winking ability is probably the result of difference in occupations and manner of life. In her occupations, woman finds less use for monocular vision, and, her life being less out-of-door, her eyes do not require lid protection so frequently.' Ten years earlier, C. Rider (1889) had reported that 30 females out of 180 could not wink, but only one male was unable to do so out of 120 examined. Danielson (1932) said 'I believe with the Riders that men wink better than women'. So far as age is concerned, Duke-Elder (1952) was of the opinion that 'the habit' (winking) 'is usually acquired in early childhood'. Danielson (1932) used similar terms, stating 'I believe from my present work that the winking ability is determined in early childhood and persists throughout life much the same as handedness'.

Changes in winking ability. In this regard, Danielson (1932) stated that: 'While I have not had a sufficiently large group for dependable data, I am quite certain that the only condition that will ordinarily ever markedly change or reverse a unilateral winking habit or modify a non-winking condition is a unilateral painful lesion associated with blepharospasm.' Duke-Elder (1952) felt that the habit 'is as difficult to change in later life as

handedness'. Schmidt-Rimpler (1889) said that: 'Some individuals find it difficult to close the one eye, and only learn it by long practice.' However, McKay (1889) cited his own experience of a spontaneous increase in winking ability.

II. Plan of investigation

The subjects. The 527 subjects of European stock were consecutive patients presenting for dental examination in Melbourne, with the exception of two, each of whom was blind in one eye, and of those who, because of the loss of maxillary incisor teeth, were not included in the main investigation, of which this study is, as it were, a by-product. The age and sex distributions are shown in Table 1A. The Polynesian sample was obtained on the island of Rarotonga, situated in the South Pacific about half-way between Tahiti and

Table 1A. *European stock*

Age in years...	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21+	Total
Male	4	15	9	13	25	13	14	23	12	19	16	11	12	9	8	3	6	2	22	236
Female	2	10	16	17	16	17	19	27	12	17	21	13	11	9	6	7	7	8	56	291
Total	6	25	25	30	41	30	33	50	24	36	37	24	23	18	14	10	13	10	78	527

Samoa. In spite of the remote situation of the island, there has been a certain amount of intermarriage with Europeans, but it is considered that the subjects included in this investigation were almost entirely of Polynesian stock. They comprised all the children attending the primary school of the most remote village on the island, all adults attending for examination at the most remote meeting house, and all the pupils of the secondary school who were 15 years of age or more. Their age and sex distribution are shown in Table 1B.

Table 1B. *Polynesians*

Age in years ...	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21+	Total
Male	3	3	6	13	10	17	6	6	6	4	7	10	12	7	3	1	3	2	12	131
Female	2	1	3	10	14	2	9	6	12	8	7	7	19	7	4	5	5	2	8	131
Total	5	4	9	23	24	19	15	12	18	12	14	17	31	14	7	6	8	4	20	262

Classification of winking ability. In this paper, by winking is meant 'to close one eye, as in aiming at a target' as defined in the *Oxford Dictionary* (1933), and not in the sense, as used by Miles (1931), of the simultaneous closing of both eyes. Six categories of eye winking were differentiated: (1) Subjects who could not close either eye without closing the other one also, in this paper termed non-winkers. (2) Those who could close only the right eye without closing the other, here called unilateral right winkers. (3) Those who could close only the left eye without closing the other, called unilateral left winkers. (4) Those who could close the right eye more easily than they could the left, termed preferred right winkers. (5) Those who could close the left eye more easily than they could the right, termed the preferred left winkers. (6) Those who could close both eyes readily and with approximately equal facility, and called in this paper the bilateral winkers. This classification is similar to that used by W. Rider (1898) in his study of visual acuity. The closing of one eye was, in most cases, accompanied by the puckering of the other. This was disregarded, although its degree varied greatly. In most cases, where difficulty was experienced in closing an eye, a considerable degree of contraction of the accessory facial muscles was seen. Little difficulty was experienced in deciding to which category each

subject should be assigned. In order to assess the error of these observations, 106 European subjects and 52 Polynesians were re-examined several weeks after their initial testing. Only one discrepancy was noted in each group, in both cases the difference being of only one winking category.

III. Results

The results are summarized in Table 2.

Table 2

	Europeans		Polynesians	
	Number	Percentage	Number	Percentage
Non-winkers	51	9·7	12	4·6
Unilateral right	50	9·5	24	9·2
Unilateral left	79	15·0	26	9·9
Preferred right	103	19·5	68	26·0
Preferred left	128	24·3	62	23·7
Bilateral winkers	116	22·0	70	26·6
	527	100·0	262	100·0

Non-winking. In those subjects who were 10 years of age and under, in both racial groups, the incidence of non-winking was significantly higher ($P < 0{\cdot}001$) than in those over 10 years old.

Bilateral winking. In those who were 10 years of age and under in the European group, the incidence of bilateral winking was significantly lower ($P < 0{\cdot}01$) than it was in those subjects who were over 10 years old. However, in the Polynesians, too many subjects developed this skill at an early age for the last-mentioned effect to be demonstrable.

Unilateral and preferred winking. Among those subjects who were unilateral or preferred winkers, in the European group 207 (57·5%) winked the left eye and 153 (42·5%) the right one ($P < 0{\cdot}01$). This preponderance of left winking is mainly due to the winking preference of those who were 10 years of age and under, of whom 100 winked the left eye and 65 the right, this difference in proportion, also, being significant ($P < 0{\cdot}01$). However, in the Polynesian group, both in the total number of subjects in these classes and in those under 10 years of age, whether the results of the unilateral and of the preferred winkers were added or considered separately, there was no significant difference between the number of subjects who winked the right eye and those closing the left one.

Winking ability and sex. Both in the European and in the Polynesian groups, no significant difference was established between the proportions of the two sexes which were seen in each of the winking categories.

Winking ability and age. In the European group a significant correlation ($P < 0{\cdot}001$) was found, in each sex, between winking ability and age, the difference between the two sexes again not proving to be significant, the pooled coefficient of correlation being 0·39. The Polynesian group gave a similar result, r being 0·26 ($P < 0{\cdot}01$). These relationships are displayed in Fig. 1, in the compiling of which the comparative winking ability was expressed by allotting the numerical value of 0 to the non-winking groups, of 1 to the unilateral winking groups, of 1·5 to the preferred winkers, who could wink one eye with ease and the other with more difficulty, and of 2 to the bilateral winking groups. The ages were grouped into 3-year periods to improve the clarity of the graphs. In Fig. 2 the upper curves display the rise in incidence of the ability to wink either eye (bilateral and preferred winkers), and the lower curves set out in Fig. 2 depict the fall in the percentage of non-winkers as age increases.

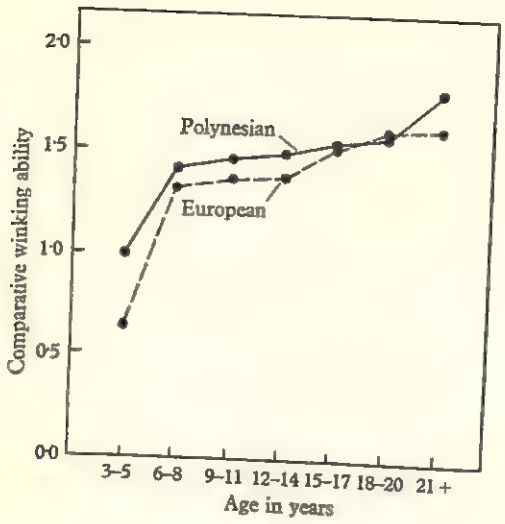


Fig. 1

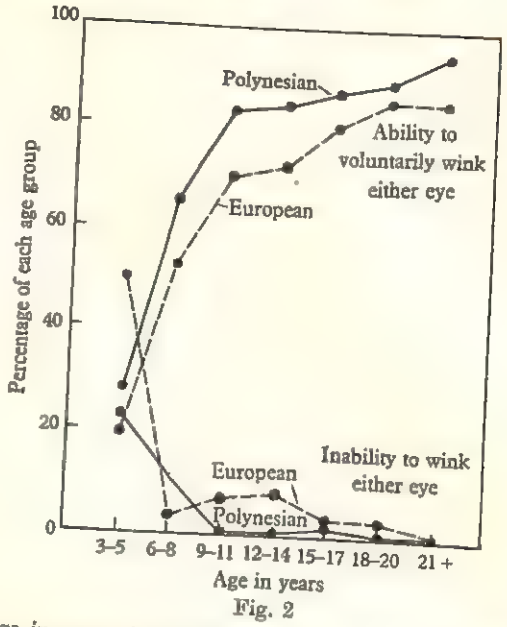


Fig. 2

Fig. 1. The relationship between winking ability and age in years in the European and the Polynesian racial groups.

Fig. 2. The relationships between the ability to voluntarily wink either eye, and age in years; and between the inability to wink either eye independently of the other, and age in years. European and Polynesian racial groups.

IV. Discussion

The incidence of right winking ability in the subjects of European stock is the same as that found by Danielson (1932) in a slightly smaller sample. However, the incidence of non-winking, 9·7%, was higher than the 7% which he reported, and that of bilateral winking (22%) was lower than his figure of 33%. In addition, the occurrence of left winking (39·3%) was higher than his finding, which was 31%. Unfortunately, Danielson did not state the ages of his subjects, but it is considered that the difference between his results and those of the subjects of European stock in the present work may be attributable to there being, in this sample, a larger proportion of subjects who were 10 years of age and less; for, in this group, there is a significantly greater proportion of those who are non-winkers and of those who wink the left eye, and a significantly smaller proportion of bilateral winkers than are to be found in the remainder of the sample.

These Polynesians spend most of their time in the open air, and their work includes few tasks which require monocular vision. Their winking pattern differs from that of the European group mainly in that their winking ability was similar on both sides, and from that group and from the data recorded by Danielson (1932) in that only 4·6% of the Polynesians were unable to wink.

The correlation between voluntary winking and age, as revealed in the present work, is of interest. when the belief of Danielson (1932) is recalled that 'winking ability is determined in early childhood and persists throughout life much the same as handedness': and that he was 'quite certain that the only condition that will ordinarily ever markedly change or reverse a unilateral winking habit or modify a non-winking condition is a unilateral painful lesion associated with blepharospasm'. Duke-Elder (1952) also was of the opinion that 'the habit' (winking) 'is usually acquired in early childhood and is as

difficult to change in later life as handedness'. The relationship between age and winking ability from 6 years to 21 years of age is probably linear, there is certainly a highly significant correlation. When the subjects under 6 years of age are included the curve is, obviously, no longer linear, and it is possible that the youngest age group forms a different population. This is suggested, also, by the similar shape of the curves obtained from the data of the two racial groups. However, the precision of the measurement is not sufficient to allow any definite statement on this point. There is still a highly significant correlation between winking ability and age; in fact, for obvious reasons, r has increased.

By excluding the non-winkers from the data, it can be demonstrated, in both races, that the shift of the point on the curve (Fig. 1) which represents the youngest age group, from the approximately linear conformation of those points which denote the winking ability of the other age groups, is due to the high incidence of non-winking in this group. This is displayed in the lower curves of Fig. 2. In this connexion the concept of C. Rider (1889) may be recalled, that there is an 'age at which monocular vision is sought by the individual . . . . I think it will be found to be at the age of about five years in boys, later in girls.' In the present study, in the European group, only 42% of those who were under 5 years old were able to voluntarily wink one or both eyes. However, this ability was seen in 80% of the group of children who were 5 years old, and it was present in over 93% by the time they were between 6 and 7 years of age. In the Polynesians, 55% of those under 5 years could wink, and the figure rose to 87% in those who were 6 years of age.

C. Rider in 1889, W. Rider in 1898, and Danielson in 1932, were all of the opinion that 'men wink better than women' (Danielson, 1932). However, in the present work, both in the Polynesians and in the subjects of European stock, the winking ability of the male group was not significantly superior to that of the female one. W. Rider (1898) said that 'this disproportion between the sexes in the matter of winking ability is probably the result of differences in occupations and manner of life'. This theory is supported by the present results, for when the European figures are compared with those presented in his paper, it is seen that the altered way of life of women during the present century has been accompanied by a significant decrease in the incidence of non-winking in those of them who are over 21 years of age. No conclusion can be drawn in regard to the younger age groups, as no figures regarding the winking ability of these groups in the past have been found.

I should like to express my thanks to Prof. Sir Arthur Amies, C.M.G., Dr J. Ringland Anderson and Dr S. B. Hammond, Senior Lecturer in Psychology, University of Melbourne, for reading the draft of this paper; and to Mr R. T. Leslie, Senior Lecturer in Mathematical Statistics, University of Melbourne, for his advice on statistical matters.

References

Danielson, R. W. (1932). Some theoretical and practical points regarding unilateral voluntary winking. *Amer. J. Ophthal.* **15**, 611–17.

Duke-Elder, W. S. (1952). *Text-Book of Ophthalmology*, Vol. v, p. 5153. London: Kimpton.

McKay, R. J. (1889). In discussion on paper by Rider, C. E. (1889).

Miles, W. R. (1931). Elevation of the eyeballs in winking. *J. Exp. Psychol.* **14**, 311–32.

Rider, C. E. (1889). The winking test. *Trans. Amer. Ophthal. Soc.* **5**, 551–59.

Rider, W. (1898). Unilateral winking as a test of comparative visual acuity. *Trans. Amer. Ophthal. Soc.* **8**, 434–52.

Schmidt-Rimpler, H. (1889). *Ophthalmology and Ophthalmoscopy*. Quoted from Rider, C. E. (1889).

(*Manuscript received* 20 *March* 1957)

SHAPE RECOGNITION: A REPLY TO DODWELL

By J. A. DEUTSCH

Institute of Experimental Psychology, University of Oxford

Dodwell (1957) criticizes the writer's theory (Deutsch, 1955) on three main grounds. First, he says that Lashley's (1938) experiments cannot be used as evidence for it, since his figures subtended too large a visual angle to be seen as a whole, while the theory postulates that they are seen as a whole. The reason, says Dodwell, why they could not be seen as a whole is that parts of them would fall outside the rat's area of distinct vision. However, Lashley's square would have subtended about 30° and is readily seen as a whole by the human eye, even though the edges be outside the area of distinct vision. We cannot form conclusions about figures being seen as a whole from considerations of distinct vision or acuity.

Secondly, he objects that rats do not treat shapes which have been rotated in the sagittal plane as equivalent. It is true that there is no good evidence for this in the rat, which has a preference for cues arising from the tilt of the figure. However, as there is no good evidence against either, and as other species appear to recognize shape as apart from tilt, it seems sound enough to frame a theory which explains this ability. In any case, the Lashley jumping stand on which the rat is tested may prejudice the issue in favour of tilt. A rat will jump to horizontal bars without training. The selection of cues of tilt has an obvious biological utility in jumping situations.

Thirdly, Dodwell objects to the writer's theory because he has found that a few of his rats can discriminate circles from squares (apparently without any training). Apart from the suspicious nature of these particular findings, Lashley did find that this discrimination presented the utmost difficulty, and Dodwell has not explained in a satisfactory manner why this should have been the case. He suggests that the brightness relationship in the various quadrants of a card containing a solid circle and a solid square are the same. Therefore, as brightness is a 'dominant' cue, the rat fails to discriminate. This line of reasoning would lead us to predict that a rat could not discriminate a square from another rotated through 45°. However, Lashley found that rats could perform this discrimination.

In connexion with the second criticism, Dodwell quotes Gellermann (1933) to show that chimpanzees do not have the ability to recognize rotated triangles, as demanded by the writer's theory. This is because they rotated their heads in the discrimination situation. But it seems that this is a misinterpretation of the evidence. It is clear from Gellermann's account that head rotation took place after recognition of the correct figure, and not before. Head turning only occurred in relation to the correct figure, even when new figures were substituted for the negative figure. 'All subjects continued to react positively to the triangle under these conditions. An interesting type of behaviour accompanied these correct responses. In every case the subjects hesitated on their trials under these new conditions, and, after looking from form to form *they turned their heads about 60 degrees to one side or the other and viewed the triangle.* Then they made the correct response.' (Gellermann's italics.) Such findings support the writer's theory, not Dodwell's.

Dodwell's rebuttal of the writer's theory of shape recognition consists mainly of calling attention to the fact that animals discriminate tilt, of which the writer has been well aware. A theory of shape recognition is unlikely to explain this capacity since a system picking out shape *per se* must discard information about tilt. If we suppose that there are two permanent lines on the two-dimensional array of cells (see Deutsch, 1955), one vertical and the other horizontal (each with its own final common cable) which send down a message when wave fronts from shapes externally imposed on the two-dimensional array arrive, then information about the tilt of lines will be available to the animal. This would solve the difficulty which Dodwell feels so much.

REFERENCES

DEUTSCH, J. A. (1955). A theory of shape recognition. *Brit. J. Psychol.* **46**, 30–7.
DODWELL, P. C. (1957). Shape recognition in rats. *Brit. J. Psychol.* **48**, 221–9.
GELLERMANN, L. W. (1933). Form discrimination in chimpanzees and two-year-old children. I. Form (triangularity) *per se*. *J. Genet. Psychol.* **42**, 3–27.
LASHLEY, K. S. (1938). The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.* **43**, 123–93.

(*Manuscript received 2 February* 1957)

PUBLICATIONS RECENTLY RECEIVED

The Measurement of Meaning. By C. E. Osgood, G. J. Suci and P. H. Tannenbaum. Illinois: University Press. 1957. Pp. 342. $7.50.

Take a sample of subjects and a sample of 'concepts', e.g. lady, boulder, sin, father: ask each subject to rate each 'concept' on a sample of 7-point bipolar scales, e.g. good-bad, robust-delicate, restless-quiet: sum the ratings of each 'concept', factor analyse, and rotate for simple structure, and you generally find the three orthogonal factors, namely, 'evaluative', 'potency', and 'activity' (some possible additional minor factors have not yet been fully worked out). Each 'concept' can be given a position in this 3-D (theoretically n-D) space, which represents its cultural meaning. Distance from the origin corresponds to intensity. This, in essence, is the method used by Osgood and his collaborators. As they point out, although every 'concept' has its own specifiable position in their semantic space, you cannot necessarily recover the concept from its three co-ordinates. To do so, you need additional information, the 'linguistic and situational context' in which the 'concept' appears. It is meaning *without context* which they claim to measure.

This seems to be a fairly obvious application of factor analysis, but one which is fraught with methodological difficulties, and which requires a good deal of spade work. It is to the credit of Osgood *et al.* that they have already accomplished so much. It is a pity that this is not more evident from a brief glance at the book; discussion and experiment are mixed together, and the reader cannot see at once which of the two any page contains. As we should expect with Osgood as senior author, the methodological difficulties are discussed frankly. For example, the authors note the failure of the seven steps on their rating scales to be of equal subjective size. They have estimated the amount of error to be expected from this failure, and have decided that it is sufficiently small to be neglected when compared with the noise level of their measuring instrument. Again, the authors note the failure of the bipolar dimension 'robust (rugged)—delicate' to pass through the zero of the evaluative factor (the whole dimension appears to be mildly 'good'). In this case they merely state that it is not possible to evaluate parts of a measuring instrument against other parts, without laying yourself open to the charge of circularity. They can as yet suggest no way of dealing with the problem.

Having frankly admitted these and other methodological difficulties, the authors proceed to ignore them, and describe some of the applications of their method. Success here is, in itself, a rough form of validation. They show that their method can give a more general description of attitudes than can be obtained by the use of the conventional single scale, which corresponds to their evaluative factor. They and their students have measured political attitudes, the conventional meanings of colours, and aesthetic judgements. They have also measured changes in attitude from exposure to information in various guises, the acquisition of meaning by novel 'concepts', and the temporary changes in meaning which occur when two incongruous concepts are associated. Roughly 60% of their more detailed predictions in this field have been borne out, and a number of novel findings have been recorded. Thus (p. 282) 'Whereas the meaning of NURSE can easily be devalued by attaching TREACHEROUS to it, attaching SINCERE to PROSTITUTE fails to budge the immoral connotation of the fallen woman.' This level of success is no mean achievement in view of our rudimentary knowledge of guiding principles in this area.

The authors and their students have also attempted to validate certain aspects of the Thematic Apperception and Rorschach Tests, and a number of hypotheses about personality and the content of dreams. Roughly 25% of the tests proved significantly successful. Finally they have made a detailed study of the three personalities in a case of triple personality, both towards the middle and end of treatment. With the minimum of background information, they made a clinical evaluation of the patient based upon their measurements. This evaluation does not appear to your reviewer to match the report of her two psychiatrists much better than by chance. However these rather unsatisfactory results need not necessarily be taken as a criticism of their method, for so many hypotheses in clinical psychology are incompatible that only a relatively small proportion are ever likely to be validated. And the reports of psychiatrists are normally loaded with subjective errors. In fact, only a relatively small amount of agreement in this clinical area is to be expected, if their approach is a valid one.

This book describes the first really successful approach to the experimental study of meaning. What conclusion do we reach? Meaning is simply three main dimensions (in a hypothetical 'semantic space')

combined with a context. At first, this may seem a little difficult to accept, given the amount of attention which meaning has received from philosophers and from some of the older psychologists. Are our doubts any different from the doubts which are always expressed by contemporaries when a longstanding mystery is explained away? I wonder.

E. C. POULTON

The Psychology of Perception. By D. W. Hamlyn. London: Routledge and Kegan Paul. 1957. Pp. vii+120. 12*s.* 6*d.*

In this book Hamlyn attempts a philosophical critique of the views of the Gestaltists, Michotte, Gibson, Hayek, and the Transactionalists on perception. His main thesis is that these writers have not been offering, as they thought, scientific explanations of perception, but have merely offered different conceptual analyses of what perception is, and that these analyses are in all cases mistaken. He maintains that a correct philosophical analysis of the concept is an essential prerequisite to any scientific explanation of perception. Unfortunately he makes no attempt to supply a correct analysis of perception himself, and the result is that the tone of the book is completely negative. It is by no means obvious that a correct philosophical analysis of a concept in ordinary usage (or rather—in the case of 'perception'—in philosophical usage) is a prerequisite to a scientific explanation of the phenomena covered by the term—Hamlyn is content with a dogmatic assertion, he does not argue the point: if the development of chemistry had had to wait upon correct philosophical analyses of words such as gold or heat, we might still be in ignorance of atomic theory to-day. What is important to the scientist is to be clear about the facts which he is trying to explain. Koffka's discussion of the question 'why do things look as they do?' is an attempt to get clear about these facts, and the useful part of his discussion is that an answer to the question must start from the physical stimulus as it affects the retina and not from the object itself.

Many of the arguments Hamlyn deploys against psychologists are put forward with such dogmatism and cursoriness that it is difficult even to make a guess at what could have led the author to present them. Thus, on page 19, he maintains that it is impossible to explain perception in terms of underlying brain processes because perception involves learning and learning cannot be explained in terms of a brain process: no reason is given for this curious belief. His main thesis against the Gestaltists seems to be that it is impossible to give explanations of normal perceptions (p. 31)—this is simply the way we perceive when there are no abnormal conditions which make us perceive abnormally. Yet any full explanation must explain both the normal and the abnormal cases. A match normally lights when struck: a full explanation will explain both why it normally lights, and why it fails to light if conditions are abnormal (e.g. if there is no oxygen in the air). What the psychologist is trying to do is to formulate the relationship between stimuli and responses and to make hypotheses about the intervening mechanisms: yet Hamlyn has succeeded in writing a book on this topic with scarcely a reference to responses and with no discussion of the term.

Hamlyn is consistently condescending and unfair to the psychologists he discusses. He makes no attempt to see what is useful about their views: e.g. it is quite true that Hayek uses the word 'classify' loosely both of the whole organism and of the nervous system. Hamlyn rightly points this out, but he does not go on to consider what Hayek was really trying to say and whether it is a useful way of looking at things.

The topics discussed in the book seem to be chosen almost at random: there is no mention of Brunswik, no discussion of explanations involving unconscious inference. Nor do some of Hamlyn's own mistakes help to inspire confidence: it is alarming to find that the author of a book entitled *The Psychology of Perception* believes that light from objects low in the visual field falls on the bottom of the retina (p. 88). Moreover, Hamlyn's own philosophical confusion seems to be at least as great as that of the psychologists he decries.

Since psychologists often do fall into conceptual confusions in discussing perception, a book written by a philosopher on the psychology of perception could have been extremely helpful to psychologists. To be helpful, however, the author would have had to make a genuine attempt to see what psychologists have been trying to say, rather than be content with trying to expose verbal confusions in their attempts to say it.

N. S. SUTHERLAND

The Scientific Study of Social Behaviour. By MICHAEL ARGYLE. London: Methuen. 1957. Pp. viii + 239. 21*s.*

Mr Argyle intends this book for students of Psychology and of the other Social Sciences. He distinguishes it from a text-book on the grounds that it deals selectively with the field and, for instance, omits discussion of socialization and of personality studies. In this book, the author wishes to put facts before theories and then to examine the various theories in the light of these facts. This procedure, Mr Argyle considers, differs from that of text-books of Social Psychology, which he describes as 'often illustrating a particular theoretic viewpoint with a congeries of experiments, quasi-experiments, and other people's opinions'. One wonders whether this description fits some of the better text-books in the field, e.g. Asch's *Social Psychology*. Mr Argyle further hopes that the book will be of use to social administrators by indicating methods of research which would help in the solution of practical problems.

It is on these grounds, then, that the book should be evaluated. It begins in a very promising way with a lively introduction in which Mr Argyle stresses that the study of social behaviour is 'a field of experimental psychology with a special subject matter'. Here, however, he limits it unduly to the study of the behaviour of people when others are present. An equally important sphere of study for the social psychologist is that of the influence on the individual of those who are not present, for example, the influences of mass media and of wider group belongingness.

The book is uneven. Some parts, like those concerned with small groups, are excellent. A wide range of research material is presented in a vivid manner and in sufficient detail to permit the reader to follow Mr Argyle's line of argument. This is not always the case in those sections in which Mr Argyle is perhaps less interested. In these sections, researches are listed but the data are presented in such a concentrated manner that it is difficult to make an assessment; also, the style tends to be more abrupt and, particularly in the first section, dealing with methods of studying social behaviour, the author does not always make it clear where description of the data ends and interpretation begins. The section dealing with the design of social investigations is well organized; in it, the author discusses a wide range of empirical designs and examines each in turn to see how far they satisfy the conditions of valid scientific inference. Chapter III, 'The Explanation of Social Behaviour', lists the various stages in theory construction. This section is very interesting. When Mr Argyle passes from the consideration of theories to generalizations, the book becomes a little too much of a catalogue. Concept after concept is described without always informing the reader of the link between them. For example, in a short section entitled 'Self Perception', we are told about accuracy of perception, about insight, about level of aspiration, about self perception, about congruence between different aspects of self perception after therapy, and then are taken back once again to level of aspiration.

The over-all impression gained is that this is an important reference book for the student of Psychology, first for literature references over an unusually wide field of social psychological problems, and second for the novel and stimulating approach of some of its sections. In particular, one would recommend those sections dealing with the design of experiments. The book has not, however, quite achieved what it set out to do. It has not integrated facts—indeed it suffers from the lack of Gestalt that theory helps to impose on data. Part of the difficulty arises from the sheer vastness of the field, part from the varying degree of the author's 'at homeness' with the material. I look forward to an intensification of Mr Argyle's obvious interest and knowledge in the study of small social groups, which should result in another book. Meanwhile, I should like to recommend his book, with the reservations mentioned, to all those who wish to familiarize themselves with recent research in the field of Social Psychology. HILDE T. HIMMELWEIT

Memory: Facts and Fallacies. By IAN M. L. HUNTER. Harmondsworth: Penguin. 1957. Pp. 185. 3*s.* 6*d.*

As most psychologists are aware there is no single volume which covers the various scientific studies of memory. Gomulicki's British Psychological Monograph represents a laudable attempt to classify the various theories which have been put forward and it is to be regretted that so far no author has attempted to do the same for the experimental data.

Dr Hunter's book, published in the Pelican Psychology Series, is naturally designed for a wide public but in many respects it is an attempt to present the experimental data; its title might well be 'Memory: Facts, not Fallacies' for Hunter is not out to expose but to present the evidence. This he has done in as straightforward, matter-of-fact an account as could be demanded. The plan of the book is to make some

appraisal of what we understand by memory, to describe the processes of memorizing and discuss the nature of forgetting and recall. This is followed by chapters on the repressing and imaging of material, and a final section is devoted to 'improving memory' where the discussion is predominantly about mnemonic systems and their relation to other efforts to classify material.

Within this framework there is no doubt that Hunter has marshalled many of the accepted experimental data of remembering. We have here a workmanlike little text and one that has studiously avoided the novel and sensational. Much recent work, particularly in the fields of short-term memory and neurology, seems to have been deliberately excluded. The most controversial chapter is that on 'Repressing' where, after indicating the influence of repression on free association, an effort is made to explain the loss of childhood experiences along conventional lines without resource to psychoanalytic interpretations. Even here the author is quick to acknowledge that we have 'too few facts and too many question marks' to decide between the processes.

Some psychologists will I think find the pedagogical tone of the book a little irksome. It is admirably direct but it can be dictatorial. Thus we are told at the beginning: 'The basic memory processes are four in number, namely, learning, remembering, forgetting, and retaining' (p. 14). This is an odd quartet, whilst those readers who think that the initial perceiving is an integral part of the process will find its omission puzzling, particularly as Hunter is keen to indicate Bartlett's contribution.

But just because this text is so unassuming and direct it does raise a major query for the psychologist. Hunter himself believes: 'The great advances of psychology in the present century have been due to the adoption of the scientific method' (p. 136). How far we consider this to be the whole truth will probably depend on what we understand as scientific method, but since this is an account for the general public of what has been accomplished in this field we, as psychologists, may well ask ourselves: 'How much has been discovered?' 'What is the nature of any advances which have been made?' 'Where do we go from here?' The point is made all the sharper by Hunter concluding his book with Fulwood's 1562 poem on 'Memorie' which prescribes amusingly close to the present-day scientific formula.

It has to be admitted that the advances are not spectacular. In the reviewer's opinion, psychologists have observed two major processes which would help us to explain the aberrations of remembering. First, our present remembering is inaccurate because in the past, due to our own individual experiences, we did not perceive and assimilate accurately (Bartlett), and secondly, something which is close to the other side of the coin, our remembering is inaccurate because a present repression is inhibiting a past percept (Freud). Evidence would suggest that Bartlett's is much the more ubiquitous experience but there seems no reason for denying the occurrence or import of the Freudian phenomena. Further advances are slight: there are several variations on the Bartlettian theme which have supported and slightly extended his original monograph, whilst the work with brain-injured amnesia patients has brought out the curious time-relations which exist between experiences. Recent studies on short-term memories have also emphasized both this and the differences between immediate and long-term recollections. But *in toto* it is a slim contribution; on the one hand the neurological picture is inevitably confused because of the difficulty of ascribing with such a complex and interacting system any one phenomenon to any particular neurological defect, whilst any general theory of remembering is inadequate because it inevitably becomes involved in the whole gamut of mental processes about which insufficient is known. I think we have to accept that future work on remembering must tackle this problem in a piecemeal and systematic fashion and it is of note that American studies in this field do attempt to do this much more than we do in Britain. It would seem to be a reminder that when we are confronted by such processes as remembering there is a need for more systematic and sustained attacks upon the problem if we are to progress.

HARRY KAY

Personality, Appearance and Speech. By T. H. Pear. London: George Allen and Unwin. 1957. Pp. 167. 21*s.*

Prof. Pear defines personality as 'the effect upon others of a living being's appearance, behaviour, etc. so far as they are interpreted as distinctive signs of that being'. A tremendous increase in contacts, direct or indirect, between people has taken place in our age of 'mass communication'. The social importance of judgements of others (based on speech, dress, mannerisms, etc.) is greater than it has ever been before; partly because of the 'effects upon others' that an individual may exercise when suitably processed by those who have recently been called the 'hidden persuaders'. What do psychologists know about all this? Very little, says Prof. Pear, and he proceeds to fill the gaps, chapter by chapter. Dress, speech, face, gesture, gait—hundreds of subtle and amusing observations follow each other. At the end of it all, the

reader is left breathless and frustrated. The book is curiously disjointed: it is as if we were invited to listen to a brilliant after-dinner conversation that Prof. Pear is having with himself. The titles of chapters give an indication of the general theme for the next few pages; what follows is a series of variations on the theme, each of them enjoyable, but only a few consecutively arranged and developed in the grim pursuit of statements of general validity.

Prof. Pear is a keen observer of the social scene. Many of the almost self-contained paragraphs are full of observations which are a challenge to further research, and will have to be taken up, sooner or later, by those interested in social perception. Neglected sources of information are often brought to light. For example, why is it that psychologists have been so little concerned with the significance of differences between national groups in their way of describing similar situations? 'The English "cope with" a difficulty, implying that success may not be either immediate or complete, Americans, more optimistically, "fix" it.' An Englishman 'stands for Parliament'; an American "runs for Congress'; An Australian says 'Too right it is' when an Englishman says 'Not half it isn't'. Or, a glimpse into a musty corner: 'Teachers in "finishing" schools and deportment classes will have lists of "do's" and "don'ts"; it would be interesting to learn how these change with time; who initiates improvements; whether they are ever resisted.' Similar indiscreet and pertinent questions are asked about beauty contests, male and female, about the ups and downs of radio, television and film 'personalities', and about innumerable other fads which are suddenly with us, as by some miracle—often carefully contrived.

Concluding his book, Prof. Pear writes: 'Omelettes cannot be made without breaking eggs: in this book several new ones have been cracked: the noises may have surprised some psychologists more than the ordinary reader.' The omelette was certainly well prepared; but this reviewer begs to disagree with the author's estimate of the number of eggs; it was not a substantial meal, and the noises were not very surprising. Prof. Pear's concern is with a problem, perennial in psychology, of the relationship between controlled investigations, conducted in the usual ways, and the rich complexity of 'real life'; he finds that there is a large discrepancy between the two. Very often this is true; the most promising way out is probably through isolating variables of higher order of complexity and applying to them, as far as possible, methods which have proved useful in simpler contexts. This is being attempted all the time, with varying degrees of success. This book will certainly stimulate some psychologists to turn their attention to some new and complex variables and, in doing this, it will perform a valuable service. But it does not go very far beyond demonstrating that complex variables do exist, and it does not contain many hints as to what should be done about it.

H. TAJFEL

A Short History of Psychotherapy. By NIGEL WALKER. London: Routledge and Kegan Paul. 1957. Pp. xii+185. 25*s*.

This book is intended mainly for non-psychotherapists with an interest in the subject. It traces the development of modern methods from Freud and Jung right down to the present day, including such topics as group psychotherapy and therapeutic communities. It is short and blissfully lacking in that semantic vagueness and exuberance which characterizes so much writing in this field. The author is not afraid to tackle such topics as the efficacy of psychotherapy, a point taken for granted by most writers. Altogether, this is an excellent book very much to be recommended to those interested in the topic. On the critical side, I found somewhat disturbing his omission of several people who have made very important contributions—more so on any reasonable appraisal than others who are mentioned. The index of names refers to Hiawatha, but not to Herzberg whose doctrine of 'active psychotherapy' was an important development. Wolberg is mentioned, but no Wolpe whose use of learning theory in connexion with psychotherapy has become one of the most important new developments in this field. However, with these exceptions, the book is remarkably complete considering the small number of pages. Presumably the author might argue that opinions and the importance of individuals differ and will continue to do so.

H. J. EYSENCK

The Sane Society. By ERICH FROMM. London: Routledge and Kegan Paul. 1956. Pp. xiv+370. 25*s*.

This book is not so much a psychological study as a tract for the times describing, somewhat after the manner of the sage of Ecclefechan, first, the parlous state into which human society has fallen, and secondly, the conditions under which humanity may escape perdition if only it will heed the prophet's

warning. The author's approach is, of course, that of the American school of culture psychologists, who, in an attempt to escape the limitations of Freud's biological approach, have directed attention so predominantly to the influence of cultural factors on individual development as to surrender the insight into unconscious motivation achieved through Freud's researches. It is a guiding principle of the author that the problem of psychological abnormality is not so much one of failure on the part of the individual to adjust to the conditions prevailing in a given society as of failure on the part of society to provide an answer to man's existential problems; and it is in the light of this principle that he feels justified in applying the terms 'sane' and 'insane' to social systems. Man's existential problem is further regarded as arising out of his having lost that harmony with 'nature' which he originally enjoyed through instinctive adaptation, and his experiencing a consequent necessity to make good the loss of this original harmony by seeking 'ever-higher forms of unity with nature, his fellow men and himself'—a necessity which is 'the source of all psychic forces which motivate men'. In the author's view, the freedom conferred upon mankind by the dissolution of his pristine harmony with 'nature' becomes a source of considerable anxiety to the individual who is, accordingly, only too liable, in his attempt to achieve security through some new harmony, to allow himself to become the puppet of the prevailing social system—whether this be (in terms of the contemporary situation) a capitalist Western society in which the individual is treated as a commodity or a Communist society in which he is treated as a slave. The only escape from the 'robotism' and consequent 'insanity' involved in both these prevailing forms of society is, in the author's view, that provided by 'Humanistic Communitarian Socialism'—which would appear to be a social system in which the individual is treated as an end rather than a means, is afforded an opportunity for creative activity, and is related to other individuals by bonds of love. The general conception of such an ideal society is, of course, not entirely novel, having been already described in the New Testament under the less cumbrous title of 'The kingdom of heaven'.

Whatever interest Fromm's theme may possess at a superficial level, his treatment of it displays the characteristic weakness of the American school of culture psychologists in that, adopting the approach of 'humanistic psycho-analysis', he seeks to describe human behaviour in terms of object-relationships while at the same time ignoring the concepts of internal reality and relationships with internal objects. His ignoring of these concepts renders it easier for him to entertain the notion of man's pristine harmony with, and subsequent loss of, 'his original home, nature' (better known perhaps as 'the garden of Eden'), while apparently remaining oblivious to the unconscious significance and motivation of the phenomena to which he refers.

W. RONALD D. FAIRBAIRN

Psychoanalysis and Psychotherapy. By FRANZ ALEXANDER. London: Allen and Unwin. 1957. Pp. xiv+299. 25*s*.

Dr Alexander, a native and medical graduate of Budapest, became a leading member of the first training institute for psychoanalysts soon after its foundation in Berlin in 1920. Since he emigrated to the U.S.A. in 1930 he has been an influential member of the American Psychoanalytical Association, particularly through his work at the Chicago Institute of Psychoanalysis and as Clinical Professor of Psychiatry at the University of Illinois. His comments upon contemporary psychoanalytic theory and practice in the U.S.A. are therefore both authoritative and interesting.

Of the essays which have been welded together into this book, the first group is concerned, broadly speaking, with the problem of explaining what takes place in the therapeutic process of psychoanalysis. The old fallacy that it consists of the extension of the patient's self-knowledge is discussed once more (perhaps this is necessary in a country where Adler's influence was so strong); and the importance of corrective emotional experience and of the patient's relationship to the analyst is stressed in a way which to British psychoanalysts will seem entirely orthodox. Dr Alexander has a distinction to make between two kinds of regression, one of which seeks an earlier stage of development because it was so satisfactory, while the other does so because the later stage was so unsatisfactory; but he does not draw any very practical conclusions from the distinction, which seems to me merely to separate two sides of the same penny. Perhaps his most practical observation is that the orthodox Freudian ideal of a wholly impersonal analyst, whose own reactions to the patient are kept to a carefully controlled minimum, is not only unattainable but also a positive hindrance to therapy; this is one of the unacknowledged ways in which post-war Freudians on both sides of the Atlantic have moved a little toward the Jungian point of view. (Its danger is, of course, that the moral judgements which orthodox Freudianism succeeded in banishing will be smuggled in again under a professional pseudonym.)

The second group of essays is concerned partly with the relationships between the pure salt of psychoanalysis and the various solutions of it which Dr Alexander labels 'psychotherapy', and partly with the problems of training practitioners of both kinds. Some of these problems are less real in this country than in the U.S.A., where psychoanalysis gained a quicker and wider acceptance in all fields of psychiatry between the wars (it has been suggested that the chief reason was the gap left by the slower development of physiological techniques of treatment). However this may be, it almost sounds as if one of the hardest tasks of psychoanalysts in the U.S.A. from now on will be, not to secure popularity for their ideas, but to maintain a pure strain from which to revitalize the many cross-bred techniques that are reproducing themselves.

NIGEL WALKER

Introductory Clinical Psychology. By S. L. GARFIELD. New York: The Macmillan Company. 1957. Pp. xiii+469. 42*s*.

The purpose of this book is to give beginners a picture of test instruments in the field of intellectual functioning and personality. In addition, it discusses the role of the clinical psychologist as research worker and therapist. There is a very general first chapter which gives what is meant to be a history of clinical psychology and a final section dealing with scientific and professional problems. Discussions of such matters as validity and reliability are not misleading because they are so general, but they do not enter into any of the complexities of the problems. This reviewer considers that such complexities might be hinted at even in an elementary text. In its discussion of tests, the book seems adequate in so far as it covers most of the recognized and widely used clinical material from America. But it shows little awareness of other material apart from the Rorschach: thus, Eysenck's work with personality tests is not mentioned at all and Shapiro's study of brain injury is ignored. This may account for the undue weight given to projection tests.

Nearly 100 pages of this more than 400-page book is given to treatment (as usual called therapy). The author then notes that there are comparatively few studies devoted to evaluating the outcome of therapy. The word 'comparatively' is ambiguous but. since the First World War, the number of such studies is by no means inconsiderable especially in view of the uniformity of their findings. It is wrong of Prof. Garfield to ignore the consistent failure of these studies to prove the value of psychological and psychiatric treatment. It is a pity that the systematic ignoring of such data on the part of clinical psychologists was not emphasized in this text. The reason, however, may be that the author proceeds to ignore the findings of the one review he mentions.

The last few chapters of the book are concerned with the American psychologist and his professional problems and are unlikely to be of much interest to clinical students here. It is hard to praise this book, but there is a relatively lively and critical chapter on research which shows that the unfortunate state of clinical research is well understood by the author just as his book exemplifies the unfortunate state of clinical practice.

N. O'CONNOR

Physiology of the Nervous System. By E. G. WALSH. London: Longmans, Green. 1957. Pp. xvi+563. 50*s*.

There has long been a need for a neurophysiology text which introduces and organizes the facts about the *inside* of the organism for those who are concerned with the behaviour of that organism. This book is welcome not only because it is aimed at psychologists, among others, but because it incorporates recent developments, both of fact and fancy, within the field of neurophysiology. In addition to the traditional headings of 'Nerve and Muscle', 'The Visual System', 'The Cerebellum,' the reader will find chapters on 'The Applied Physiology of Pain', 'The Sensory Pathway, Thalamus and Parietal Lobe, Speech', with frequent reference to material with which the psychologist will be comfortably familiar—such as the tilted room experiments of Witkin, the inverted field experiments of Stratton, the descriptions of postural development by Gesell. The book is profusely illustrated, sometimes to the extent of containing figures for which no reference can be found in the text. The psychologist is introduced to topics of current excitement and controversy in neurophysiology—the brain stem reticular formation (although this receives less attention than it would appear to merit), the 'visceral brain', the efferent control of sensitivity of muscle spindle receptors, the work of Weddell and co-workers on somatic perception. The author presents relevant applications of modern servo theory to some of the subject matter. Frequent and intelligent reference is made to neurological case material.

It is not by way of adverse criticism to note that the reader will not find in this book a *systematic* treatment of neurophysiology. It is more a collection of interesting facts and speculations emitted, so it would appear, as they occurred to the author. Thus, we shift abruptly from the structure of the insect eye to the unidirectionality of synaptic conduction. Even in the chapter headings there is no attempt to impose a rigid organization.

Since the book is 'intended to fill the gap between the physiologist on the one hand and the clinician and psychologist on the other', questions arise which apply to all gap-fillers. Can a gap be filled when the boundaries on either side are not very firm? Assuming firm boundaries, how safe is the bridge connecting them? On both questions the book leaves something to be desired. The author's attempt to be 'psychological' is laudable, but often his psychology is not as firmly established as even that of the introductory student of the subject, who would not confuse 'instrumental' or 'operant' learning with Pavlovian conditioning, nor depend frequently on concepts such as 'sensorium', nor make unqualified assertions such as: 'The human being uses many visual cues for orientation, the rabbit none' (p. 156). Or again 'In the human being it seems that what can and cannot be done [posturally] has to be discovered by *experiencing in infancy*' (pp. 161–2, italics in original). Or 'Conditioning is most readily achieved if there is a simple relationship between stimulus and reinforcement. These principles are perhaps the most enduring feature of the work on conditioned reflexes' (p. 221). It is fruitless to argue who has the greater responsibility to learn the other chap's subject—the psychologist or the neurophysiologist. Nevertheless, it is the case that psychology can (and in the opinion of many, such as Skinner, should) exist without reference to neurophysiology, while neurophysiology (and more especially neurology), in so far as it attempts to relate the behaviour of parts of the inside of the organism to the behaviour of the whole organism, cannot escape at least a description of such behaviour. Psychology has, over the past twenty years, developed both descriptive systems and behavioural theories which could prove useful to those in related disciplines who wish to talk of behaviour.

It is also clear that there is a large gap between the gap-fillers in neurophysiology and the gap-fillers in psychology—even when they are studying precisely the same phenomena. The present book—or any other neurophysiology text—really does not close that gap. Thus, while the author has discovered Lashley, he does not—nor do most physiologists—read the journals in which psychologists customarily publish their experimental contributions to neurophysiology (notably the *Journal of Comparative and Physiological Psychology*). Thus, in discussing the frontal lobe syndrome there is no mention of several important papers by Harlow, Settlage, Pribram, Mishkin, and others. Equally important work on the effect of temporal neocortical lesions and limbic system lesions receives no mention. The extensive work of Beach on the cortical control of sexual behaviour is not noted. This list could be extended and elaborated indefinitely. In each case a reference to material in psychological journals would have contributed *directly* to topics already discussed in the text and, more important, would sometimes have altered the argument significantly. At the same time this book makes a much greater effort to bridge the gap than do most neurophysiology texts, and it is safe to predict that it will be referred to frequently by students and specialists alike.

LAWRENCE WEISKRANTZ

Extrasensory Perception. By G. E. W. Wolstenholme and E. C. P. Miller (eds.). London: Churchill. 1956. Pp. ix+240. 27*s.* 6*d.*

The papers and discussions which make up this notable Ciba foundation symposium of 1955 offer material to suit many tastes. Most of the twenty-three symposiasts live in Britain, but visitors from France, Holland, Spain and the U.S.A. also took part. Many hold scientific posts, some are medically trained and at least one is a philosopher. Partly because so much of the book is taken up with the discussions, it will be more profitable to give the impressions made by the book than to present something like a table of contents. The foremost impression is of the variety of attitudes to extrasensory perception which can be adopted by men of parts, all willing to listen to what the other man has to say. One finds, for instance, the attitude that the physicists must now rewrite their systems to accommodate the facts of para-psychology, a rewriting analogous perhaps to that forced by the revelation of the difficulties inherent in the concept of simultaneity. Then there is the view that physics need not be rewritten, but that physiology must admit some hitherto unrecognized sensory mode. Others again require no rewriting of theories; it is by training, deliberate or otherwise, that skill in particular directions is attained by individuals, equal to that of the best conjurers, who can, however, be relied on to simulate or reproduce practically any phenomena yet observed. Another view is that there is nothing to explain and that the

extraordinariness of the phenomena is (like beauty) in the eye of the beholder. A fifth is that science need not be rewritten to accommodate the phenomena, since they are different in kind from physical events and do not proceed within the physical domain; and finally some participants would, I think, adopt the position that odd and interesting things happen, which we do not yet understand but which will in due course be resolved by the application of scientific method.

When all is said and done, I can discover no hint that any member of the conference modified his attitude in any degree as a result of what was said. No comment on Dr Wasserman's 'fields' suggests to me that his hearers had understood him. Pigeons may or may not be clairvoyant, but Dr Matthews seems at the end to be no less in the dark as to how clairvoyance could be of service to get them home; and, for all Dr Parkes's optimism, many a glass of sherry will need to be downed before the difference between Dr Amadou and Mr Spencer Brown is dissipated.

Finally, there remains to puzzle us the contrast between the complexity and meaninglessness of the controlled situations in card guessing experiments, on the one hand, which leave the layman baffled, sceptical and unconvinced and, on the other, the spontaneous phenomena, so pregnant, so unforeseen, and so seldom repeated as to arouse in the scientist the worst suspicions. The former type of phenomenon has become a battleground for intellects; it is from the latter that the study of extrasensory perception renews its inspiration and draws its sustenance.

B. BABINGTON SMITH

Some Recent Developments of Psychology in Great Britain. By F. C. BARTLETT. Istanbul University. 1957. Pp. viii+92. 2·65 T.L.

In 1956 Sir Frederic Bartlett gave a series of six lectures at the Faculty of Letters, Istanbul University, which documented some of the major developments in British (one might almost say in Cambridge) psychology over the past forty years and highlighted the social changes which brought them about. Two further lectures were delivered at the Technical University of Istanbul and repeated in Ankara to largely military audiences. All eight of these lucidly brief lectures are printed here, first in Turkish and then in English. The lectures are admirably suited to their purpose of giving a clear picture of what an important cross-section of British psychologists have been doing in recent years, of how they have set about doing it, and of why. Any historian of the present century's psychology would be unwise not to read this little book.

I. M. L. HUNTER

Ciba Foundation Colloquia on Ageing. Vol. 3: *Methodology of the Study of Ageing.* Ed. by G. E. W. WOLSTENHOLME and C. M. O'CONNOR. London: Churchill. 1957. Pp. x+202. 32*s.* 6*d.*

In 1954 the Trustees of the Ciba Foundation decided upon special measures for the encouragement of laboratory and clinical investigations relative to the problems of ageing. Three conferences—or 'colloquia'—have been organized with this particular aim in view. This volume contains the proceedings of the third colloquium in this series. The contents range from a paper on 'The Study of the Ageing of Cells', by J. F. Danielli, to 'Twin Data on the Genetics of Ageing', by Franz J. Kallmann. Three of the twelve papers are likely to be of special interest to psychologists.

F. Verzár, of the Physiological Institute at the University of Basle, offered an interesting emphasis on 'adaptation' as the most suitable basis for gerontological research. This seems absolutely sound, provided one sticks closely to the definition of adaptation as 'the general capacity of living organisms to live under continuously changing conditions', and this reviewer is in close accord with Verzár when he states: 'When this adaptability decreases, ageing begins.'

Irving Lorge, of Columbia University, New York, presents a competent survey of the methodological problems faced by the psychologist in the study of intelligence and emotion in ageing. One is startled, however, to read that 'intelligent behaviour may be conceived as the capacity to make good, adequate and acceptable responses'.

The third paper is by A. T. Welford, of Cambridge, and is a model of its kind. He deals with methodology in the study of changes in human performance, and in so doing draws attention to point after point the importance of which it is difficult to exaggerate. It is to be hoped that every psychologist newcomer to the investigation of changes with age will have this paper drawn firmly to his attention.

Although the book includes the full discussion on each paper—sixteen people were present in addition to those giving invited papers—this but serves to remind us that the indefatigable organizers of these

top-level occasions have a very hard task indeed. This said, let us express the sincere hope that they will go on trying: the potentialities of the small colloquium appear far greater than do those of some large congresses.

ALASTAIR HERON

Human Relations in Interracial Housing. By D. M. WILNER, R. P. WALKLEY and S. W. COOK. Minnesota University Press. London: Cumberlege. 1956. Pp. xiv+167. 32*s*.

This book is an attempt to assess the relative influence of living in integrated or segregated housing projects on the attitudes and behaviour of white Americans. The research was carried out on two pairs of housing projects which were as alike as possible except that one in each pair had an integrated, and the other a segregated pattern of living. Within each area the relations between proximity, and the variables of contact, social climate and ethnic attitudes were studied. As there were only 10% of negro residents in each project the concepts of 'far' and 'near' were applicable within each separate housing estate and, thus, an added check was obtained. In general, it was found that contact with negroes did influence the attitudes of the white people towards them. 'Whether we consider the initially more favorable or initially less favorable respondents, those who live near Negroes in a project are more likely than those living farther away to report neighborly contact, to anticipate that white friends in the project will approve of such contact, to have high esteem for the Negroes in the project, to approve of the biracial aspect of the project, and to have a favorable attitude to Negroes in general. These "near-far" differences may be interpreted as reflecting change in attitude at least for "nears". It is of interest to note that such changes take place even among those with an initially less favorable attitude.'

This book is a valuable contribution to the growing volume of work in social psychology done on people as they are in a real-life context. In addition, the writers have tried to apply the rigorous methods of experimental analysis. It should be of interest both to people working on race relations and group dynamics, as well as to people engaged in administering mixed racial groups.

MADELINE KERR

Modern Sociological Theory. Ed. by H. BECKER and A. BOSKOFF. New York: The Dryden Press. 1957. Pp. xiv+756. $6.50.

The index of this huge book shows that Max Weber is referred to on 107 of the 760 pages of text, together with other early European sociologists like Durkheim, von Wiese and Gurvitch. American sociologists who were born or trained on the Continent also receive frequent mention—Becker (161 pages), Sorokin and Znaniecki, together with Talcott Parsons who has translated and interpreted the early writers. A glance at the list of twenty-six contributors shows that none of them is at what are generally considered to be the main centres of sociology in America—Chicago, Columbia, Michigan and Harvard. All this suggests that the book is not representative of the main stream of current American sociology, but of that minority group whose members look back to pre-1914 European thought. They would not approve of the above statistics, since they prefer more informal data which their theories help us to understand, not to predict. They might approve the following description of what we may call the 'Theorist' type of sociologist, which we shall contrast with the 'Predictor' type of sociologist with whom we are more familiar. (*a*) Theorists try to gain some understanding of phenomena by dividing them up in illuminating ways into different types of groups, leaders, etc. (*b*) They are critical of 'empiricism' on the grounds that it produces mere collections of facts, leads to unimportant research, and diverts attention from the important theoretical issues (pp. 190ff.); they prefer suggestive, though less precise information, such as can be obtained from case studies and historical research. (*c*) They are interested in the sociology of knowledge, law, religion, art and (for some reason) peasants. Predictors have a taste for cities, industry, small groups and public opinion. Both are concerned with social class and the family. (*d*) Theorists are preoccupied with logical and metaphysical matters on which empirical data have no bearing; they are concerned with the historical origins of their ideas and they write in a scholarly and verbose manner.

The development of such typologies, or classifications, is important, and it does add to our understanding. However, the theorist takes the problem no further—the enlightening classification is his final goal. He does not go on to inquire, for example, what the conditions are which make a sociologist a Theorist rather than a Predictor. For the Predictor, such classifications are an essential first stage, he then proceeds to find the empirical laws describing functional relations between variables so described.

However, there are certain empirical questions involved in constructing typologies which theorists do not face squarely: How high, for instance, are the correlations between the various elements in the typology? Not all the contributors to the present book are opposed to orthodox empirical methods—to continue our analogy. Predictors generally proceed from typologies to dimensions, and have empirical methods for deciding which set of dimensions is the best.

What is sociological theory anyway? The theorists regard their classifications as theories. Predictors are mainly interested in theories as a source of predictions to be tested—and no predictions follow from the Theorists' theories. There is little in sociology to correspond to the theories in psychology or the natural sciences—partly because there are so few laws to explain, partly because no one is agreed on what sociological theories would be like, and because neither Theorists nor Predictors are interested in deductive theories. One possibility is explanation by reduction to psychological findings, and a few miniature systems of this sort have been constructed, mainly in the small group area.

The first part of the book, 'Major Strands in Theory and Methodology', expounds the Theorist position for 300 pages, with much historical reference: the present reviewer did not find it very illuminating. The following section 'Some Specialisations in Modern Sociology' has a good chapter on religion by Honigsheim, together with the expected ones on knowledge, law and, surprisingly, small groups. The dead hand of the Theorist falls more heavily on some of these chapters than others. 'Convergences of Bordering Fields with Sociology' contains an interesting chapter on psychoanalysis by Hinkle. The final section 'Sociological Research and Theory Abroad' has an interesting account of British sociology by Professor Sprott, and also covers France, Germany, Italy and Japan; this is the most interesting part of the book.

The book is beautifully produced, contains a wealth of material on a variety of subjects, has extensive bibliographies and a good index, and is reasonably priced.

MICHAEL ARGYLE

How Children Learn: an Educational Psychology. By A. N. FRANDSEN. New York and London: McGraw-Hill. 1957. Pp. xiv+546. 41*s.* 6*d.*

This is a workmanlike text by an author who is familiar with both the academic literature and the school-room and who never loses sight of his subject-matter, namely, the real-life child in process of learning. In contrast to several recent texts, it avoids excess of, on the one hand, arid ivory-tower theory and, on the other, banal superficial anecdote. The bulk of the book is devoted to the relations between learning proactivity and the seven factors of maturity, teacher-guidance, practice, perception of the effects of provisional tries, transfer, motivation, and emotional disturbances. Then follow discussions on appraising individual differences in abilities, providing for these differences, and appraising school achievement. For the British student undergoing teacher training, the text's usefulness may be limited by its American setting. It should, however, be of great assistance in providing orientation and content to those concerned with giving courses in either educational psychology or the psychology of learning.

I. M. L. HUNTER

Sense and Nonsense in Psychology. By H. J. EYSENCK. London: Penguin Books. 1957. Pp. 349. 3*s.* 6*d.*

This is good light reading and will appeal widely by reason of its style, which is lucid and intelligible, and its subject-matter, which is frankly popular—hypnosis, lie detectors, telepathy, dreams, personality, politics, aesthetics. Prof. Eysenck is always interesting, often entertaining, occasionally irritating. When he gets on to psychoanalysis his critical level sadly deteriorates—any kind of tendentious argument or trivial anecdote will do to belabour his *bête noire*. This is a pity: every sensible psychologist is heartily sick of this tiresome guerrilla warfare; and there is no need for it, since psychoanalysts lay themselves open in plenty to serious and responsible criticism. Fundamentally I believe Eysenck is much too confident about things. There is not perhaps quite such a sharp and certain division between sense and nonsense as he would make out. It is somewhat surprising to be told that the main outline of the picture of human personality supplied by Pavlovian theory is unlikely to require any revision! Are the foundations of this theory so very certain? In the history of science the sense of yesterday not infrequently becomes tomorrow's nonsense; and nonsense turns out after all to be sense. I suspect the future development of psychology holds some surprises in store; and perhaps some major revisions of outlook after all! But breezy confidence makes good popular fare!

L. S. HEARNSHAW

Alfred Adler: Apostle of Freedom. By PHYLLIS BOTTOME. (3rd ed.). London: Faber and Faber. 1957. Pp. 300. 25*s.*

Adler's stock is rising, so the third edition of Phyllis Bottome's biography, first published in 1939, has appeared at an opportune time. The author, a novelist, was an ardent disciple and close personal friend of Adler, and she gives a vivid account of his life and personality. To psychologists and students, however, her book will be of less value than Lewis Way's *Alfred Adler*, published recently in the Pelican Psychology Series.

MARGARET KNIGHT

Symbols of Transformation. Volume Five of the Collected Works of C. G. Jung. Ed. by HERBERT READ, MICHAEL FORDHAM and GERHARD ADLER; transl. by R. C. F. HULL. London: Routledge and Kegan Paul. 1956. Pp. xxix+567. 35*s.*

The Analysis of Dreams. By MEDARD BOSS, transl. by ARNOLD J. POMERANS. London: Rider, 1957. Pp. 223. 25*s.*

Mysticism, Sacred and Profane. By R. C. ZAEHNER. Oxford: The Clarendon Press, 1957. Pp. xviii+256. 42*s.*

Followers of Jung will be glad to see the fifth volume of this admirably produced series. The book is lavishly illustrated, with 64 plates and 43 text-figures, and it is sufficiently well known to need no further comment.

Prof. Boss's Analysis of Dreams contains an interesting Foreword by Dr E. B. Strauss, and is an attempt to approach dreams in a phenomenological and existentialist way. The author has studied Freudian and Jungian interpretations of dreams, and become discontented with both. After criticizing these and other theories of dream-interpretation he puts forward his own approach.

Prof. Zaehner explains that the purpose of his book on mysticism is to solve, perhaps presumptuously, the problem of the proper relationship between the monist's felt conviction that soul and God are one thing on the one hand, and the direct confrontation of 'I' and 'Thou' on the other. He attempts to distinguish between what seem to be radically different types of mysticism and to relate them to one another.

R. W. PICKFORD

The Child and the Family. By D. W. WINNICOTT. London: Tavistock. 1957. Pp. ix+147. 12*s.* 6*d.*

The Child and the Outside World. By D. W. WINNICOTT. London: Tavistock. 1957. Pp. viii+190. 16*s.*

These books contain a collection of papers written or broadcast during the past 18 years by Dr Winnicott. The papers' arrangement provides some continuity of ideas but does not prevent some overlapping and repetition which becomes a little tedious if the books are read at a sitting. The first volume is addressed to the parents of young children but is likely to be read and appreciated more by those professionally concerned with the welfare of children since, as the author suggests, the ordinary mother is usually too much involved in practical tasks to reflect much on theoretical aspects of family relationships. The first section is designed to give confidence to the 'ordinary devoted mother' who wishes to be allowed to get to know her baby and to begin feeding and caring for it in her own way, without constant interference from doctors and nurses. It provides a simple account of psychoanalytical views on the mother-child relationship, infant feeding, and the child's early emotional and intellectual development with the minimum of reference to the origins of these ideas. The more intelligent lay parent might well inquire what evidence exists for some of the statements. Some readers may consider the author's views of the 'ordinary devoted mother' and the natural wisdom she brings to her task as being over-optimistic but, in other respects, many of his practical implications—on, for example, feeding, weaning, and toilet training—are now fairly generally accepted by paediatricians.

The second volume is addressed to teachers and social workers concerned with the care of children not their own. One section deals with nursery school education, sex education, adoption and related problems: another with 'Some reflexions on impulse in children'. The third section concerns the problems of evacuated children during the recent war and contains important implications for the problems of separation and children's institutions. Dr Winnicott was closely concerned with the setting up of hostels for difficult evacuated children and his account of the problems encountered and the methods for dealing

with them provides much information of theoretical and practical value today. One talk on 'The deprived mother' draws attention to a side of the separation problem which is frequently neglected. While this third section is perhaps the most interesting for the psychological reader, the two books together provide a worthy response to the requests of a wider audience for the publication of Dr Winnicott's works. His style is lucid, his manner friendly, and his years of experience provide much wise insight into child behaviour and parental attitudes even for those who do not accept his theoretical bias.

AGNES E. NAUGHTON

Psychopharmacology. Ed. by NATHAN S. KLINE. Washington: American Association for the Advancement of Science. 1956. Pp. 165. 32*s.*

Techniques for the Study of Behavioral Effects of Drugs. By P. B. DEWS, B. F. SKINNER, and ten others. 1956. *Annals of the New York Academy of Sciences*, **65**, 247–356. 21*s.*

The Pharmacology of Psychotomimetic and Psychotherapeutic Drugs. By S. S. KETY, and fifty-five others. 1957. *Annals of the New York Academy of Sciences*, **66**, 417–840. 35*s.*

In the last few years 'psychopharmacology' has become a subject for busy research, especially in the United States. New 'tranquillizing' and 'hallucinogenic' drugs have attracted the attention of many different kinds of investigators, neurophysiologists, biochemists and psychiatrists, as well as psychologists and pharmacologists. Co-operation between them is growing because the relations between the central nervous system and the drugs which act upon it are extremely complex.

Psychopharmacology deals, in spite of its general title, almost entirely with the clinical uses of chlorpromazine and reserpine. These drugs are now widely given to make psychiatric patients more manageable. They are believed to do this by calming them without making them drowsy or inefficient, but case histories and clinical trials show that individual differences are large. The last three papers examine modes of action of several drugs, including mescaline. Lysergic acid, the 'psychotic' effects of which were discovered by accident in 1943, produces profound psychological disturbances in doses as small as 0·0001 milligram.

Techniques for the Study of Behavioral Effects of Drugs is mainly about animals, especially operant conditioning. The effects of drugs and of reinforcement schedules may mimic each other; for example, effects of a stimulant drug on the pecking rate of pigeons may resemble those of periods of *ad libitum* feeding. Operant behaviour can discriminate reliably between drugs, and this could be useful for screening new substances. In the only paper on human subjects, interesting methods are described for analysing moods. Different drugs produced characteristic moods, and group experiments showed that moods due to some drugs were more contagious than those due to others.

The Pharmacology of Psychotomimetic and Psychotherapeutic Drugs consists of 35 papers, mostly reviews and many of them highly specialized. 'A review of comparative behavioral pharmacology' refers to over 150 titles. The reviewer of 'Pharmacotherapy in psychiatry' finds that opinions of the value of drugs in psychotherapy are divided and finally declares himself 'conservatively enthusiastic' about chlorpromazine.

Several of the papers deal with lysergic acid diethylamide (LSD) and serotonin. Serotonin causes isolated smooth muscle to contract, and this effect is reduced by LSD. Since serotonin occurs naturally in the brain, this suggested that LSD might produce its artificial psychoses by interfering with it. Serotonin might thus have a role in maintaining normal mental functions, and be involved in real-life psychoses. If a normal mouse is placed facing downhill on a tiled surface it pushes backwards with its forelegs, as if to prevent itself sliding. LSD makes mice walk backwards, and this has been attributed to an 'hallucination' of sliding downhill. Injections of serotonin sometimes protect them from this symptom. Antagonism between the two substances can also be shown with other responses, such as conditioned avoidance behaviour. In other experiments, however, both on behaviour and with isolated tissues, the effects of serotonin and of LSD have been not opposite but similar. Moreover, several drugs antagonize serotonin in isolated tissues but have no psychological effects.

The common aim of all this work is to help explain mechanisms of the central nervous system and of mental illness. The findings reported are new and interesting. It is not surprising that they leave many questions unanswered, especially when so little is still known about drugs like alcohol and barbiturates which have been studied for many years. 'I feel strongly', says Prof. Kety summing up, 'that these data are trying to tell us a story, and I am happy that so many of us appear to be listening.'

HANNAH STEINBERG

OTHER PUBLICATIONS RECEIVED

Being and Nothingness: an Essay on Phenomenological Ontology. By JEAN-PAUL SARTRE. Transl. and introduced by H. E. BARNES. London: Methuen. 1957. Pp. lxvii + 636. 50*s.*

The Tibetan Book of the Dead. Third ed. By W. Y. EVANS-WENTZ. With a psychological commentary by C. G. JUNG. London: Oxford University Press. 1957. Pp. lxxxiv + 249. 25*s.*

The Road to Inner Freedom: the Ethics. By B. SPINOZA. Edited and introduced by D. D. RUNES. New York: Philosophical Library. 1957. Pp. 215. $3.00.

Personality Traits between Puberty and Adolescence: their Relationships, Development and Constancy with Reference to their Relation to School Achievement. By R. O. VIITAMÄKI. Helsinki: Finnish Academy of Science and Letters. 1956. Pp. 183. F. mk. 700.

The Art of Marriage. By MARY MACAULAY. Harmondsworth: Penguin. 1957. Pp. 127. 2*s.* 6*d.*

Problems of Human Pleasure and Behaviour. By MICHAEL BALINT. London: Hogarth Press. 1957. Pp. 300. 25*s.*

Hamlet's Mouse Trap: a Psychoanalytic Study of the Drama. By ARTHUR WORMHOUDT. New York: Philosophical Library. 1956. Pp. 221. $3.50.

Truth and Meaning. By DAVID GREENWOOD. New York: Philosophical Library. 1957. Pp. 114. $3.75.

Experimental Psychology and Other Essays. By I. P. PAVLOV. New York: Philosophical Library. 1957. Pp. 653. $7.50.

The Nature of Conflict: Studies on the Sociological Aspects of International Tensions. By The International Sociological Association, in collaboration with J. BERNARD, T. H. PEAR, R. ARON and R. C. ANGELL. Paris: UNESCO. London: H.M.S.O. 1957. Pp. 314. 21*s.*

Neurotic Interaction in Marriage. Ed. by VICTOR W. EISENSTEIN. London: Tavistock Publications. 1957. Pp. 352. 30*s.*

The Psychopathology of Prostitution. Rev. ed. By EDWARD GLOVER. London: Institute for the Study and Treatment of Delinquency. 1957. Pp. 20. 1*s.* 6*d.*

The Problem of Homosexuality. Ed. by EDWARD GLOVER. London: Institute for the Study and Treatment of Delinquency. 1957. Pp. 40. 3*s.*

Family and Kinship in East London. By MICHAEL YOUNG and PETER WILLMOTT. London: Routledge and Kegan Paul. 1957. Pp. 232. 25*s.*

Patients and Doctors. By KENNETH WALKER. Harmondsworth: Penguin. 1957. Pp. 182. 3*s.* 6*d.*

Culture and Mental Disorders. By RALPH LINTON. Illinois: Thomas. Oxford: Blackwell. 1956. Pp. 139. 34*s.*

The Mind Goes Forth. By HARRY and BONARO OVERSTREET. London: Peter Davies. 1957. Pp. 350. 21*s.*

Psychiatric Studies. By C. G. JUNG. Volume One of Jung's Collected Works. Transl. by R. F. C. HULL. London: Routledge and Kegan Paul. 1957. Pp. 269. 25*s.*

Methods and Criteria of Reasoning: an Inquiry into the Structure of Controversy. By RUPERT CRAWSHAY-WILLIAMS. London: Routledge and Kegan Paul. 1957. Pp. 296. 32*s.*

Envy and Gratitude: a Study of Unconscious Sources. By MELANIE KLEIN. London: Tavistock Publications. 1957. Pp. 101. 12*s.* 6*d.*

Schizophrenia: Somatic Aspects. Ed. by DEREK RICHTER. London: Pergamon Press. 1957. Pp. 181. 40*s.*

Individual Differences in Night-Vision Efficiency. M.R.C. Special Report No. 294. By M. H. PIRENNE, F. H. C. MARRIOTT, and E. F. O'DOHERTY. London: H.M.S.O. 1957. Pp. 83. 8*s.*

The Stress of Life. By HANS SELYE. London: Longmans, Green. 1957. Pp. 324. 18*s.*

Selected Contributions to Psycho-analysis. By JOHN RICKMAN. Ed. by W. C. M. SCOTT. London: Hogarth Press. 1957. Pp. 411. 30*s.*

Disaster: a Psychological Essay. By MARTHA WOLFENSTEIN. London: Routledge and Kegan Paul. 1957. Pp. 231. 23*s.*

Psychological Services for Schools. Ed. by W. D. WALL. Hamburg: UNESCO Institute for Education. 1956. Pp. 150. 5*s.*

The Trickster: a Study in American Indian Mythology. By PAUL RADIN. London: Routledge and Kegan Paul. 1956. Pp. 211. 21*s.*

Changing Concepts of Psychoanalytic Medicine. Ed. by S. RADO and G. E. DANIELS. New York and London: Grune and Stratton. 1956. Pp. 248. $6.75.

Experimental Design and its Statistical Basis. By D. J. FINNEY. London: Cambridge University Press. 1955. Pp. 169. 30*s.*

Man and Automation. By L. LANDON GOODMAN. Harmondsworth: Penguin. 1957. Pp. 286. 3*s.* 6*d.*

Consumer Problems and Personal Finance. Third ed. By ARCH. W. TROELSTRUP. London and New York: McGraw-Hill. 1957. Pp. 511. 45*s.*

The Psychology of Sexual Emotion: The Basis of Selective Attraction. By VERNON W. GRANT. London: Longmans, Green. 1957. Pp. 270. 35*s.*

Structural Psychology: De Humani Mentis Fabrica. By D. and K. STANLEY-JONES. Bristol: Wright. 1957. Pp. 179. 21*s.*

Nostalgia: a Psychoanalytic Study of Marcel Proust. By MILTON L. MILLER. London: Gollancz. 1957. Pp. 306. 21*s.*

The Psychology of Persuasive Speech. Second ed. By ROBERT T. OLIVER. London: Longmans, Green. 1957. Pp. 466. 35*s.*

Probability and Scientific Inference. By G. SPENCER BROWN. London: Longmans, Green. 1957. Pp. 154. 15*s.*

The Placement of Adoptive Children. By J. RICHARD WITTENBORN. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 189. 36*s.*

Mental Depressions and their Treatment. By SAMUEL HENRY KRAINES. New York and London: The Macmillan Company. 1957. Pp. 555. 56*s.*

Applicability of Factor Analysis in the Behavioral Sciences: a Methodological Study. By STEN HENRYSSON. Stockholm Studies in Educational Psychology, No. 1. Stockholm: Almquist and Wiksell. 1957. Pp. 156. Sw. kr. 15.00.

Studies in Hysteria. By JOSEF BREUER and SIGMUND FREUD. Translated and edited by J. and A. STRACHEY. London: Hogarth Press. 1957. Pp. 335. 25*s.*

Essays in Linguistics. By JOSEPH H. GREENBERG. Chicago University Press. London: Cambridge University Press. 1957. Pp. 108. 22*s.* 6*d.*

Alcoholism: Basic Aspects and Treatment. Ed. by HAROLD E. HIMWICH. Washington: American Association for the Advancement of Science. London: Bailey Bros. and Swinfen. 1957. Pp. 212. 52*s.*

Measurement and Evaluation in Education. By J. M. BRADFIELD and H. S. MOREDOCK. New York and London: The Macmillan Co. 1957. Pp. 509. 38*s.* 6*d.*

Sociometric Testing: A Guide for Teachers. By M. L. NORTHWAY and L. WELD. Toronto University Press. London: Oxford University Press. 1957. Pp. 72. 16*s.*

Cultural Pluralism and the American Idea: An Essay in Social Philosophy. By H. M. KALLEN. Pennsylvania University Press. London: Oxford University Press. 1957. Pp. 208. 40*s.*

Philosophical Psychology. By J. F. DONCEEL. London: Sheed and Ward. 1957. Pp. 363. 21*s.*

The Meaning of Persons. By P. TOURNIER. Transl. from French by E. HUDSON. London: S.C.M. Press. 1957. Pp. 238. 21*s.*

The Psychologic Study of Man. By J. MONEY. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 216. 36*s.*

The Mentally Ill Child: A Guide for Parents. By S. B. GETZ and E. L. REES. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 88. 25*s.*

The Three Faces of Eve. By C. H. THIGPEN and H. M. CLECKLEY. London: Secker and Warburg. 1957. Pp. 313. 18*s.*

The Challenge of Soviet Education: The Study of Education as a Weapon. By G. S. COUNTS. London: McGraw-Hill. 1957. Pp. 330. 45*s.*

Man and Culture: An Evaluation of the Work of Malinowski. Ed. by RAYMOND FIRTH. London: Routledge and Kegan Paul. 1957. Pp. 292. 32*s.*

Statistical Theory: The Relationship of Probability, Credibility, and Error. By LANCELOT HOGBEN. London: George Allen and Unwin. 1957. Pp. 510. 45*s.*

Soviet Education. Ed. by G. L. KLINE. London: Routledge and Kegan Paul. 1957. Pp. 192. 21*s.*

Executive Performance and Leadership. By C. L. SHARTLE. London: Staples. 1957. Pp. 302. 30*s.*

Living Magic: The Realities Underlying the Psychical Practices and Beliefs of Australian Aborigines. By RONALD ROSE. London: Chatto and Windus. 1957. Pp. 222. 18*s.*

The Dynamics of Interviewing: Theory, Technique, and Cases. By R. L. KAHN and C. F. CANNELL. New York: Wiley. London: Chapman and Hall. 1957. Pp. 368. 62*s.*

Pain and Pleasure: A Study of Bodily Feelings. By THOMAS S. SZASZ. London: Tavistock Publications. 1957. Pp. 301. 30*s.*

Light, Colour and Vision. By YVES LE GRAND. Transl. from French by R. W. G. HUNT, J. W. T. WALSH and F. R. W. HUNT. London: Chapman and Hall. 1957. Pp. 512. 63*s.*

Mental Health Aspects of Urbanization. By several authors. London: World Federation for Mental Health. London: H. K. Lewis. 1957. Pp. 46. 6*s.*

Mental Health and the World Community. Ed. by FRASER BROCKINGTON. London: World Federation for Mental Health. London: H. K. Lewis. 1957. Pp. 118. 10*s.*

Sigmund Freud: Life and Work. Vol. III. *The Last Phase, 1919–1939.* By ERNEST JONES. London: Hogarth Press. 1957. Pp. 536. 35*s.*

How Children Learn to Speak. By M. M. Lewis. London: Harrap. 1957. Pp. 143. 8*s.* 6*d.*

Identity and Interpersonal Competence. By N. N. Foote and L. S. Cottrell, Jr. Chicago University Press. London: Cambridge University Press. 1956. Pp. 305. 37*s.* 6*d.*

Closed Ranks: an Experiment in Mental Health Education. By E. Cumming and J. Cumming. Harvard University Press (Commonwealth Fund). London: Oxford University Press. 1957. Pp. 192. 28*s.*

Problemkinder. By Various Authors. Bremen: Institut für Jugendkunde. 1957. Pp. 64. DM. 4.20.

The Feeling of Superiority and Anxiety-Superior: the Ottawa Trial Survey on Personality. By E. L. Remits. Ottawa: Runge Press. 1957. Pp. 82. No price given.

Peter and Caroline: A Child Asks about Childbirth and Sex. By S. Hegeler. Transl. from Danish by M. Michael. London: Tavistock. 1957. Pp. 30. 7*s.* 6*d.*

Helping Yourself with Psychiatry. By Frank S. Caprio. New York: Prentice-Hall. London: Bailey Bros. and Swinfen. 1957. Pp. 256. 40*s.*

The Magic of Believing for Young People. By Claude Bristol. New York: Prentice-Hall. London: Bailey Bros. and Swinfen. 1957. Pp. 204. 24*s.*

On the Unified Factor Theory of Mind. By Y. Ahmavaara. Helsinki: Finnish Academy of Science and Letters. 1957. Pp. 176. F.mk. 750.

Psychotherapy: A Modern Theory and Practice. By E. L. Phillips. London and New York: Staples. 1957. Pp. 334. 30*s.*

Homicide in American Fiction, 1798–1860: A Study in Social Values. By D. B. Davis. Cornell University Press. London: Oxford University Press. 1957. Pp. 346. 30*s.*

Systematic Sociology: An Introduction to the Study of Sociology. By Karl Mannheim. Ed. by J. S. Erös and W. A. C. Stewart. London: Routledge and Kegan Paul. 1957. Pp. xxx+169. 24*s.*

The Educational Attainment of Bilingual Children in Relation to their Intelligence and Linguistic Background. By W. R. Jones, J. R. Morrison, J. Rogers and H. Saer. Cardiff: University of Wales Press. 1957. Pp. 51. 3*s.* 6*d.*

The Doctor Eyes the Poor Reader. By D. G. Schubert. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 101. 27*s.* 6*d.*

Poor Monkey: The Child in Literature. By Peter Coveney. London: Rockliff. 1957. Pp. 297. 30*s.*

New Developments in Analytical Psychology. By Michael Fordham. London: Routledge and Kegan Paul. 1957. Pp. 214. 25*s.*

An Introduction to Social Anthropology. Vol. 2. By Ralph Piddington. Edinburgh: Oliver and Boyd. 1957. Pp. 819. 30*s.*

Psychology and Psychological Medicine for Nurses. By Portia Holman. London: Heinemann. 1957. Pp. 156. 10*s.* 6*d.*

Glory Reflected: Sigmund Freud—Man and Father. By Martin Freud. London: Angus and Robertson. 1957. Pp. 218. 21*s.*

Hypnosis in Health and Sickness. By G. Ambrose and G. Newbold. London: Staples. Pp. 196. 15*s.*

Gruppen Psychotherapie. Ed. by H. Hiltmann, K. H. Wewetzer and H. R. Teirich. Bern and Stuttgart: Hans Huber. 1957. Pp. 331. DM. 16.

PARTIAL ADVANCE INFORMATION IN A CHOICE REACTION TASK

BY J. ALFRED LEONARD

Medical Research Council's Applied Psychology Research Unit, Cambridge

An experiment was carried out to study the effect of partial advance information on performance. Reaction time and accuracy measures were compared for three conditions: a six-choice task, a three-choice task, and a set of experimental conditions in which a two-choice task preceded two three-choice tasks by various durations of time. Tests under all conditions were carried out after varying amounts of practice. It was found that at all stages of practice a difference between performance on the six-choice task and on the three-choice task was maintained. As the separation between the two-choice and three-choice task was varied from 0·04 sec. to 0·50 sec. in the experimental conditions, performance became increasingly similar to, and finally identical with, that under the control three-choice condition. The relationship between duration of forewarning and performance was, however, not a simple one. The results are discussed with reference to traditional and current views on information processing.

I. INTRODUCTION

The experiment to be described stemmed originally from work on anticipation in sensory-motor skills (Leonard, 1953, 1954), and was given an additional impetus by Hick's (1952) paper on the rate of gain of information. In discussing the relationship between 'receptor-effector' anticipation and anticipation based on past experience, I stated that the latter was frequently not entirely adequate to prepare responses, because 'very often there is a residue of uncertainty, a moment to moment randomness in the environment'. Receptor-effector anticipation enables one to reduce this residue of uncertainty just before a point of necessary action is reached, so that performance does not merely become faster but also smoother, responses merging one into the other without hesitation or raggedness. It was possible to show this in two experimental situations in which performance was compared when a new signal was given either at, or before, a choice point was reached in serial choice reaction tasks. This signal was thought of as providing the final item of information necessary to complete a given response, the rest of the information being provided by instructions and the subject's eventual experience of the task; in other words, anticipation based on past experience. From those experiments it seemed only a natural step to study the effects of providing the final item not in one but in several stages, to study the effects of partial advance information. The influence of Hick's work arises from the fact that in order to account for his finding of a constant rate of gain of information, he suggested as a reasonable possibility an internal process of progressive binary classification of information (1952, p. 22). It will be seen that the experimental situation chosen for the study of partial advance information was essentially one in which information could be presented to the subject in a progressively classified manner.

A method of doing this in a serial task comparable to those used in the earlier experiments would have been desirable from the point of view of continuity of research. But since practically nothing was known about the effects of partial advance information in skilled performance it seemed wiser to study these first in a setting free from possibly masking serial effects. As my experimental situation I took, therefore, a straightforward

six-choice task. It is possible to think of this as a two-choice task followed by a three-choice task. Essentially the experiment was a comparison of three conditions: a six-choice task, a three-choice task, and a set of conditions in which a two-choice preceded a three-choice task by varying amounts of time.

Now the last set of conditions can be thought of as one in which information is presented in a progressively classified manner. It was just such a system of *internal* progressive classification which Hick suggested as a possible mechanism to account for his finding of a constant rate of gain of information. There were two main questions to be answered: first, how far does the two-choice signal have to precede the three-choice signal for a subject to be able to respond to the latter as efficiently as in a pure three-choice situation? And secondly, what is the relationship between performance when all the information necessary is presented simultaneously and when it is presented successively, i.e. between the pure six-choice task, and the progressively classified task?

II. Method and procedure

Subject. The data to be presented were those obtained by myself acting as the subject.

Apparatus. The display consisted of six small neon bulbs stuck through a blackened piece of metal, so as to form an inverted V, with three bulbs to each arm. The display panel was shielded to reduce glare and was placed about 3·5 ft. from *S*, far enough to allow all bulbs to be seen easily when looking at the centre of the V. The control consisted of six micro-switches with paxolin flaps, raised at an angle of about 50°. The hands rested on an inclined panel so that the keys had to be pressed in the direction of the palm, i.e. a fairly natural clenching movement. Circuits were so constructed that the apparatus could deliver sequences of 48 stimuli in any order desired.

Conditions. The unit of presentation was a trial in which one had to respond to 48 signals. Each new signal appeared 1·5 sec. after a correct response had been made. There were two *control* conditions, namely six-choice and three-choice; and a group of *experimental* conditions in which a two-choice preceded a three-choice by varying amounts of time.

In a trial under the *six-choice condition* all the six lamps would light up together. After 0·1 sec. one of the six lights would go out.

In a trial under the *three-choice condition* the three lamps of only *one* set came on together, this set, i.e. left or right, being the same one throughout the trial. After an interval of 0·1 sec. one of the three lights went out.

In a trial under the *experimental condition* the lights of only one set came on together at any one time, but it could be either of the two sets. After an interval which was constant for any given trial, but which could be varied between trials, one light went out.

The coming on of the lights will be referred to as the *warning signal*, going out of one light as the *specific signal*, the time between as the duration of *forewarning*.

It will be seen that under the control conditions the warning signal indicated that a specific signal would have to be responded to in 0·1 sec. In the experimental condition the warning signal indicated the one of the two sets of three lights in which a specific signal would occur after a known interval. It was by this means that information was presented in two stages.

Ideally it would have been desirable to include trials in which the control conditions were given under all durations of forewarning. This would have added considerably to the times for both the running of the experiment and analysis of data. The 0·1 sec. interval for the two control conditions was chosen as a standard after extensive preliminary trials had shown that varying the duration of the forewarning in the control conditions had no facilitating effect. If anything, durations beyond 0·3 sec. tended to retard performance. (See Klemmer (1956) for recent evidence.)

Under all conditions the *correct response* was to press the key spatially corresponding to the light which had gone out. This resulted in putting out the remaining lights of the display. After an interval of 1·5 sec. the next warning signal came on. *Pressing a wrong key* caused a buzzer to sound; there was no effect on the visual display. All the keys could be pressed simultaneously so that if a correct and a wrong key were pressed together, the light went out and the buzzer sounded.

Recording. This was carried out with N. T. Welford's SETAR (1952). Time readings were taken to the nearest 0·01 sec. In all conditions, except that of the experimental condition with the shortest forewarning, the intervals recorded were as follows: (1) between onset of warning signal and onset of specific signal (forewarning); (2) between onset of specific signal and subject's first response (reaction time); (3) between subject's response and onset of warning signal (rest interval).

Sequences. The same basic sequence of 48 signals constituting one trial was used throughout the major part of the experiment. This sequence had been specially constructed with the following properties: (1) there were an equal number of signals for each category of choice, i.e. in any trial each light would appear an equal number of times; (2) runs of more than two identical signals did not appear (this applied to the sequence when used for the six-choice and the experimental conditions only); (3) diagram frequencies were balanced as far as possible. The basic sequence is reproduced below:

125663445326433614315514654242116636213532524152

Six six-way single pole switches enabled the sequence to be permuted when required, so that the subject need not have the same succession of signals twice, while the basic properties of the sequence were maintained. For the three-choice condition the sequence was transformed by always putting two of the permuting switches on the same position. In the earlier stages of the experiment a specially designed three-choice sequence had been used but the later method had been shown to produce the same results, and saved a great deal of time.

Circuits. The main components were electro-magnetic relays of various kinds. The forewarning and rest-interval periods were condenser controlled. The motivating device (see below) had an electronic timing circuit. All the apparatus—except, on some tests, the counter attached to the motivating device —was outside the experimental room.

Scoring. Two scores were obtained: duration and accuracy of each response. The time score taken was the interval between the onset of the specific signal and the subject's first response to it, regardless of whether this was a correct or an error response. It should be noted that since all the main data to be presented are averages obtained from all six fingers under all conditions finger differences are taken care of. The same applies to possible variations among keys and relays. Times were averaged for sets of trials, a trial consisting of 48 responses. Error scores were obtained directly from the SETAR record. Rates of gain were calculated in the manner described by Hick (1952, pp. 25–6). It should be noted that there are two ways of doing this for the experimental conditions: one can think of them either as six-choice tasks with the information given in two stages; or one can think of them as three-choice tasks only, ignoring the warning stage. In the former case one has to consider an input of 2·59 bits and divide the transmitted information by the sum of the forewarning period plus the reaction time observed; in the latter case one treats the hands separately, considers an input of 1·58 bits to each hand, and divides the information transmitted by the reaction time only. Treating the hands separately is perhaps only justified as long as there are no confusions between hands; since this obtained and the reaction times observed for the longer durations of forewarning were closer to the three-choice than to the six-choice times, the rates of gain for the experimental conditions with all but the shortest duration of forewarning were so calculated.

Motivation. Knowledge of results could be provided by means of an all or none indicator. The subject could activate an electromagnetic counter provided he responded correctly below a given but adjustable time limit. The counter could be either in the experimental room or outside with the rest of the apparatus.

Procedure. The over-all plan was to carry out tests at various stages during extended practice. A total of at least 63,000 responses under the various conditions was made during the 8 months of experimentation. The procedure during practice and between test series was not altogether uniform, so that one is perhaps not quite justified in presenting the data in the form of 'learning curves'. The main differences among the procedures were in the per cent of errors tolerated, the order in which the conditions were given, and the different uses of the motivating device. Of these the first is by far the most important and should be clearly stated: for the *first* test (June, after about 9600 responses) I tried to make as few errors as possible but did not reject any trials with high error incidence. In the *second* test (September, after about 26,400 responses) I tried to keep the number of errors constant for all conditions and rejected trials having more than three errors. In the *third* test (November, after about 52,800 responses) I paid no attention to the number of errors made at all. This last test had the most responses and was the most rigorously conducted of the three. Conditions were given in ascending and descending series, i.e. from six-choice through the various experimental conditions to three-choice where left and right hand were used alternately on succeeding trials. Trials were only rejected on the grounds of apparatus failure: four in 168. There was a 3 min. interval between trials, and 15 min. between each series of seven trials. On the

first two days of the third test two ascending and two descending series were given, and on the remaining five experimental days four such series each, a total of 24 complete series, 12 ascending and 12 descending. The error indicator was not used. The motivating device's counter was in the experimental room and the difficulty so adjusted that about 80–90% of the responses produced a click. It should be noted that the level was adjusted according to the condition used.

III. Results

Distributions of reaction times for three conditions of the September test are presented in Fig. 1. They are fairly typical of all distributions obtained. Since error level was constant and low throughout that test a comparison of these distributions is perhaps more meaningful than those obtained at the other tests. They indicate that the use of the

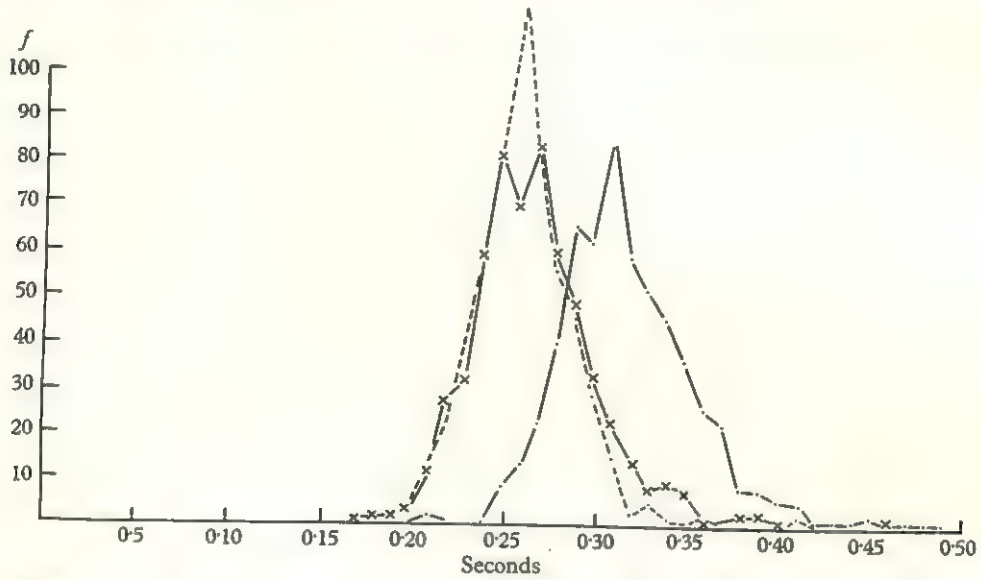


Fig. 1. Reaction time distributions for the September test for the two control conditions and the experimental conditions with 0·5 sec. forewarning. ·——·, six-choice, $N = 576$, $m = 0{\cdot}319$; ·– – –·, Exp. 0·50, $N = 576$, $m = 0{\cdot}262$; ×——×, three-choice, $N = 576$, $m = 0{\cdot}268$.

arithmetic mean of reaction times can be considered as justified. Fig. 1 shows the separation of the distributions for the two control conditions. Table 1 shows the results in terms of numbers of responses per condition, mean reaction time, and number of errors, and information measures, where calculated (see scoring above). The values for the experimental and three-choice conditions are averages for both hands.

Fig. 2 shows reaction time and error percentage curves for the three tests; the horizontal axis is of course an arbitrary scale, with the values for six-choice on the left, followed by those for the experimental conditions, and those for the three-choice on the right-hand side. This seemed a convenient way of presenting the three sets of data graphically; continuity between six-choice and three-choice is at this time an open question. It will be seen that very much the same pattern of results was obtained in all three tests.

(1) There was a clear difference between performance under the two control conditions.

(2) In the experimental condition there was a measurable effect on reaction time when the duration for forewarning was varied in the sense that broadly speaking as this duration

was increased reaction time became shorter: (*a*) the largest single effect of reducing the reaction time was obtained with 0·1 sec. forewarning; (*b*) there was some retardation of the effect with forewarning of 0·2 sec.; and (*c*) as the duration of forewarning reached 0·5 sec. reaction time tended to be equal to or better than the three-choice reaction time.

(3) Concerning the information measures it should be noted that: (*a*) on all three tests the rates of gain obtained for the six-choice control condition were higher than those for

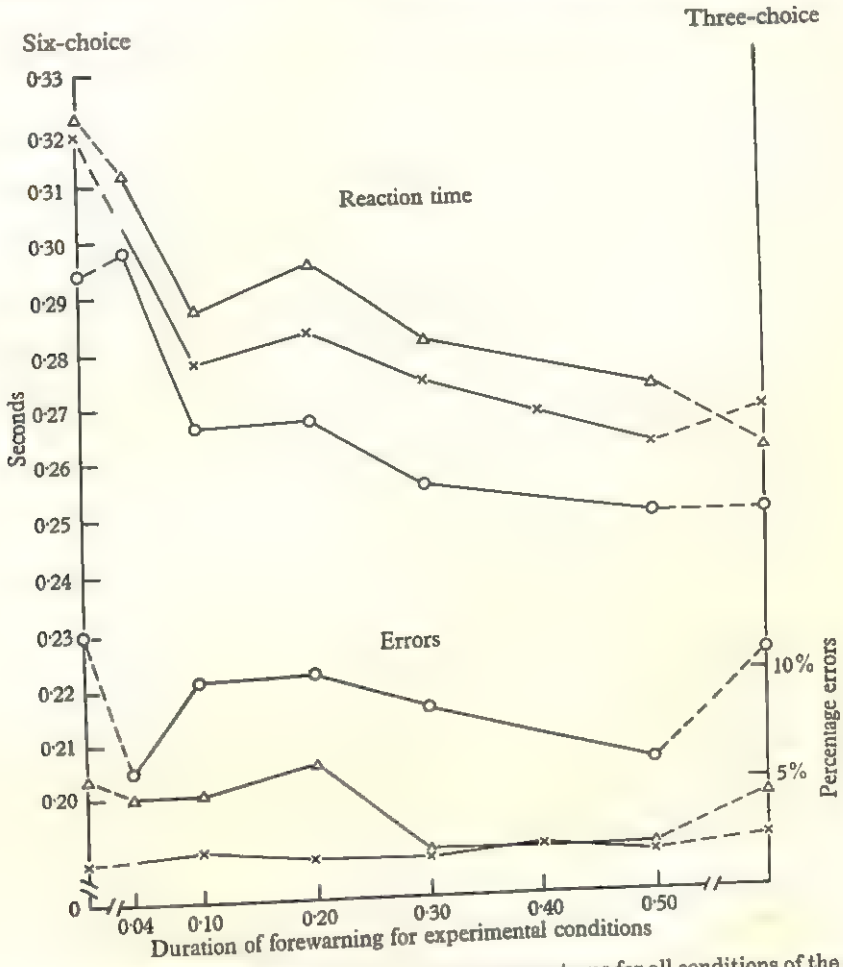


Fig. 2. Curves showing mean reaction times and error percentages for all conditions of the three tests: △——△, June; ×——×, September; ○——○, December.

the three-choice conditions; (*b*) on the last test the rate of gain for the experimental condition with 0·04 sec. forewarning came nearer to that of the six-choice condition; and (*c*) on all three tests rates of gain for the experimental condition with 0·3 sec. forewarning were comparable to those of the three-choice condition, while at 0·5 sec. forewarning rates of gain were higher on all three occasions.

Practice effects. In view of the fact that conditions were not entirely constant throughout the whole course of the experiment, progressive effects among identical conditions on different tests cannot be ascribed solely to practice. Table 2 is a reorganization of some of the data shown in Table 1 to show how reaction times, errors, and rates of gain changed

for the two control conditions on successive tests. Rate of gain, which takes into account both speed and accuracy of performance, is here a useful measure of efficiency. The highest rates of gain were obtained for the September Test. Reaction times were lowest on the December Test, but not low enough to compensate for the higher error level, so that rates of gain were lower than in September.

Table 1. *Results of three tests taken at various times*

(For methods of calculating rates of gain see scoring above)

Condition ...	Control Six-choice	Experimental (Duration of forewarning in sec.)						Control Three-choice
		0·04	0·10	0·20	0·30	0·40	0·50	
					June			
Responses	575	575	576	576	576	—	576	1152
R.T. (sec.)	0·323	0·312	0·287	0·295	0·281	—	0·273	0·261
Errors	35	29	29	36	12	—	12	49
H_T (bits)	2·24	2·27	—	—	1·44	—	1·43	1·29
Rate of gain (bits/sec.)	7·80	7·21	—	—	6·73	—	6·90	6·85
					September			
Responses	576	—	576	528	576	576	576	576
R.T. (sec.)	0·319	—	0·278	0·283	0·274	0·268	0·262	0·268
Errors	11	—	13	9	10	12	11	14
H_T (bits)	2·44	—	—	—	1·46	—	1·47	1·44
Rate of gain (bits/sec.)	8·41	—	—	—	6·96	—	7·31	7·06
					December			
Responses	1104	1102	1152	1104	1152	—	1103	1152
R.T. (sec.)	0·294	0·298	0·266	0·267	0·255	—	0·249	0·249
Errors	140	67	119	116	101	—	67	123
H_T (bits)	1·91	2·26	—	—	1·11	—	1·14	1·02
Rate of gain (bits/sec.)	7·68	7·50	—	—	6·52	—	6·75	6·44

Table 2. *Time, error, and rate of gain scores for the two control conditions only*

		June	September	December
Six-choice	R.T. (sec.)	0·323	0·319	0·294
	Errors (%)	6·1	1·9	12·7
	Rate of gain (bits/sec.)	7·80	8·41	7·68
Three-choice	R.T. (sec.)	0·261	0·268	0·249
	Errors (%)	4·3	2·4	10·7
	Rate of gain (bits/sec.)	6·85	7·06	6·44

IV. Discussion

Before answering the two questions posed at the end of the Introduction it is necessary to draw attention to the major 'incidental' finding: there was, throughout, a clear difference in the results obtained under the two control conditions. The difference in reaction time was rather smaller than that expected (e.g. Woodworth, p. 333, or Hick, p. 15) and, conversely, that in rate of gain was unexpected (Hick, 1952). While this paper was in preparation, a considerable amount of evidence appeared to show that the slopes describing the relationship between entropy and choice time are in turn related to the particular combinations of input and output coding used in a given experiment (for a discussion see Crossman, 1956). The present results show that for a particular input-output coding variation in entropy continued to show an effect even after a considerable amount of practice.

The answer to the first question (namely, how far does the two-choice signal have to precede the three-choice signal for a subject to be able to respond to the latter as efficiently as in a pure three-choice situation?) is best given by reference to the rate of gain score: the best estimate is 0·30 sec. Rather surprisingly, efficiency improved consistently when the separation was further increased up to 0·50 sec. It is interesting to note that May (1917), using the method of controlled association, also had a maximum effect of fore-warning when the duration was 0·50 sec. (for a brief description see Woodworth, 1938, pp. 791–3).

The second question concerned the relationship between performance when all the information necessary is presented simultaneously and when it is presented successively. Using the rate of gain score again, it would appear that the measures are comparable only with the shortest duration of forewarning used. (It will be recalled that to make this kind of comparison one has to add the duration of forewarning to the observed reaction time and consider an input of 2·59 bits. Since the observed reductions in reaction time due to increases in the duration of forewarning are small in relation to that duration the rates of gain will become increasingly smaller.) It is worth noting that the shortest duration of forewarning it was possible to achieve with the apparatus available turned out to be close to the observed absolute difference between the reaction times for the two control conditions. It would clearly be of some interest to establish whether the above-mentioned comparability of rates of gain was purely fortuitous or whether it is always found when the duration of forewarning is similar to that of the absolute difference in reaction times, regardless of the absolute values of the observed reaction times for two control conditions.

A second incidental result was the consistent retardation of the facilitating effect of forewarning when its duration was 0·20 sec. Similar results have since been presented by Elithorn (personal communication, 1956). I interpret this retardation as a possible instance of refractoriness (for a discussion see A. T. Welford, 1952). The point of interest here is that whereas in classical refractoriness experiments a pair of stimuli has to be responded to by a pair of responses, in the present case only a single overt response was required. The findings suggest that refractoriness may not be due exclusively to the making of a first response.

Future work on partial advance information in this area will have to take into account the effects of differences in possible input-output relationships. It may not be necessary to go to the same length in training but to concentrate instead on the three durations of forewarning mentioned above.

Grateful acknowledgements are due to Mr A. T. Welford, Director of the Nuffield Research Unit into Problems of Ageing, for advice and facilities generously offered; to Dr N. T. Welford for services psychological and electronic; to Dr W. E. Hick for stimulation and reinforcement. This work was carried out while the author was in receipt of a stipend from the Medical Research Council and the Kenneth Craik Award for 1952–3. It was contained in A.P.U. 217/54.

References

Crossman, E. R. F. W. (1956). The information capacity of the human operator in symbolic and non-symbolic control processes. In *The Application of Information Theory to Human Operator Problems.* London: M.O.S. Report WR(D) 2/56.

Hick, W. E. (1952). On the rate of gain of information. *Quart. J. Exp. Psychol.* **4**, 11–26.

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *J. Exp. Psychol.* **51**, 179–84.

Leonard, J. A. (1953). Advance information in sensorimotor skills. *Quart. J. Exp. Psychol.* **5**, 141–9.

Leonard, J. A. (1954). An experiment with occasional false information. *Quart. J. Exp. Psychol.* **6**, 79–85.

May, M. A. (1917). The mechanism of controlled association. *Arch. Psychol.* **39**.

Welford, A. T. (1952). The 'psychological refractory period' and the timing of highspeed performance, a review and a theory. *Brit. J. Psychol.* **43**, 2–19.

Welford, N. T. (1952). The electronic digital recording machine, the SETAR. *J. Sci. Instrum.* **29**, 1–4.

Woodworth, R. S. (1938). *Experimental Psychology.* New York: Holt.

(*Manuscript received* 5 *November* 1956)

FACTORS AFFECTING INDUCTIVE PREDICTIONS

By R. T. GREEN

University College, London

Two guessing tasks which appear highly similar in nature nevertheless evoke opposite trends in predictive behaviour. Attempts to generate both sets of results from the same simple principles fail because of the different implicit assumptions the subject makes about the nature of the task. It would seem that one of the assumptions that the subject tends to make when required to employ inductive reasoning, at least in the laboratory situation, is that a certain amount of stimulus variability is to be expected.

The theoretical implications are that, far from being contradictory or mutually exclusive, the gestalt and behaviourist accounts of the learning process are complementary. Only by recognizing that some aspects of the learning process are continuous while other aspects are discrete and that both S–R learning at an unverbalized level and conscious hypothesis testing go on side by side, is it possible to make sense of the more complex type of learning situation where a temporal pattern is involved.

An attempt to analyse pre-solution guessing behaviour tends to favour a non-continuous hypothesis, but in certain respects the question is still open to a more sophisticated interpretation.

I. Introduction

Learning theory has long been sharply divided over such questions as the role of reinforcement, and whether the S–R or the S–S paradigm meets the facts. More recently, we have seen these paths converging as a process of mutual adaptation makes the different approaches complementary instead of contradictory. On the surface the continuity-discontinuity controversy appeared to be an instance of outright warfare between the gestalt hypothesis-testing, insightful account of the learning process, and the behaviourist explanation in terms of gradual learning on a trial-by-trial reinforcement basis. Yet, after ploughing through the literature pro and con, experiment and counter-experiment, one is left with the distinct impression that both parties are right. In particular, consider the work of Lawrence (1949, 1950), who found that his rats learned to attend to the *relevant* cues in a discrimination task as demanded by gestalt theory, yet conformed to behaviourist principles in other respects. Prentice (1949), sympathetic to the gestalt view, found that his human subjects learned to discriminate in a way which baffled both continuous and non-continuous types of interpretation. And the earlier behaviourist approach of Spence (1940) runs into difficulties when more subtle discriminations based on transverse patterning are involved, as Bitterman (1953) points out.

Two main factors manage to obscure a great deal which the two viewpoints have in common. There is, first, a tendency for both to say the same thing in different ways. Hull (1952) ended up by leaning so heavily on concepts like the 'pure stimulus act' and 'fractional antedating goal reaction' that 'insight' became indistinguishable from 'implicit trial and error' in any operational sense. The second factor that has helped to give an unjustified impression of incompatibility is that, whereas the behaviourists have based their position on the statistical trends of groups of subjects, the gestalt theory is more concerned to describe what the individual subject is doing. Nowhere is this more apparent than when we employ human subjects for a concept-formation task using the accepted paradigm of reversal in the pre-solution period. The data of Walk (1952) seem to lend

support to the behaviourist contention; the more pre-solution practice that is given on the reversed discrimination, the more difficult it is to establish the final concept. Yet no one disputes the fact that human subjects test hypotheses while searching for a solution. All the behaviourist can do is to deny the relevance of this fact to the actual learning that occurs. At this juncture, the gestalt theorist can turn the tables very neatly: he can show that the result is quite compatible with his theory and, simultaneously, he can convict the behaviourist of an inherent logical contradiction. Accepting that the subject tries out hypotheses, then, the longer the pre-solution period on the reversed discrimination, the greater the probability that the subject will test what is to *become* the correct hypothesis and abandon it. As long as he remains unaware that the task has been changed he is more likely to avoid the correct hypothesis under the new conditions. Thus, the experimental evidence is in no way at variance with the gestalt position. Now for the logical contradiction. It is an accepted condition of this type of 'crucial' experiment that the original learning on the reversed discrimination must not go beyond the pre-solution period (see McClelland, 1943). Yet according to the behaviourist account in terms of a gradual, trial-by-trial acquisition of 'habit strength' there can be no such thing as a pre-solution period. From trial *1* to trial *999* there is no point at which the problem can be said to be solved or unsolved. Habit strength increases logarithmically to an asymptote. Learning is never perfect nor utterly imperfect. Attempts to establish the pre-solution period by showing that scores are not above chance either demonstrate that the gradual acquisition of habit strength theory is wrong or that the statistics are insensitive to the learning that must have occurred. Spence (1952) may protest that behaviourist learning theory was developed to deal with the behaviour of rats and is therefore not proven or disproven by what human subjects do. But behaviourists have only themselves to blame for such a misunderstanding, since no one has been more eager than they to explain human behaviour in S–R terms on every possible occasion. Whenever they are criticized as 'rat' psychologists, their defence is that they are seeking basic principles using lower organisms in the hope of developing a theory to deal with the more complex behaviour shown by human subjects. Spence (1940) is not averse to citing his own experiments on chimpanzees in this context.

One is driven to conclude that both schools are describing, in their own terms and for the most part accurately, different aspects of the same process. There is little doubt that the subject tests hypotheses until he finds one that pays off, or runs out of hypotheses. He continually abandons and modifies these hypotheses in the light of success and failure. He may learn that a certain type of hypothesis tends to pay off better than any other and then concentrate his search in this area. In behaviourist terms, he is learning to attend (respond) to the relevant stimuli and ignore the irrelevant. It is precisely this attending to the relevant stimuli that the gestalt theorists insist is essential for learning to take place. At the same time they overlook the possibility that, although the subject attends to the relevant stimuli, he may be directed to do so by the wrong hypothesis. In which case, it would be quite reasonable to expect some learning to take place in the S–R reinforcement manner although the problem remains unsolved. Errors for the group may therefore tend to show a steady decline. Again, if a subject hits the correct hypothesis, his errors will cease but group results will smooth out this 'insight'. In any case, we never know what process of implicit trial and error leads up to, and hence makes less dramatic, this sudden individual drop in overt errors. Furthermore, as Hull (1920) showed, it is

possible for a subject to respond 100% correctly and yet be unable to verbalize the concept.

All this preamble has been intended to emphasize the need for a more eclectic approach to the learning process if we are to understand some of the more subtle and complex phenomena that keep intruding themselves into the data and upsetting our highly polished theories. Polish is admirable—on timber that has the quality to take it. The experiments to be described pose knotty problems unless one is willing to adopt an open-minded view about what goes on when people learn; that is, modify their responses in the light of past experience. One is never sure just how far back the experience goes which is helping to determine a particular response: one does know that it extends far beyond the beginning of any laboratory experiment. To demonstrate the importance of this influence the binary guessing behaviour of a group of subjects who have to predict a light sequence is contrasted with that of another group who are faced with a very similar binary sequence using verbal material. Another light sequence reveals something of the way a simple alternation habit is learned.

II. Experimental procedure

Method 1. *Light guessing.* The apparatus consisted simply of two 3·5 V. bulbs operated independently by finger keys and supplied by a small dry cell. The light on the right of the subjects was backed by red paper and referred to as the 'red light', while that on the left was designated 'yellow' in a similar fashion.

The subjects, 121 in all, were students at a technical college, aged 15–35 with a mean age of about 23 years. Six groups were involved, two on each of three series to counterbalance for position or colour preferences. The structures of the three series (shown in Table 1) followed those of a previous memory experiment involving numbers and nonsense syllables. It had been hoped in this way to interpret differential recall probabilities in terms of unexpected structural change, but an unexpected variable proved to be more in need of interpretation.

Table 1. *Structure of the three light-guessing series*

	Serial position									
Groups	1	2	3	4	5	6	7	8	9	10	11	12
I	X	O	X	O	X	O	X	O	X	O	X	O
II	X	X	X	O	O	O	O	O	O	X	X	X
III	X	X	X	O	X	X	X	O	O	O	O	O

X, red or yellow light; O, light of alternative colour

Formal instructions to the subjects, run in groups of ten or twenty, were as follows. 'This is an experiment to find out something about the way people work out what is coming next. As you can see, there are two small lamps facing you, one with a red and the other with a yellow patch around it. One or other of these will light up briefly after each trial. Your task is to predict whether it will be the red or the yellow side which will light up. There is no guarantee that they will light up an equal number of times—one or other of them may not light up at all. All you can be sure of is that, after you have made your prediction, either the red or yellow side will light up. So now, against figure "one" at the top of the left-hand column of your slip of paper write either an "*R*" or a "*Y*" according as to whether you expect the red or the yellow side to light up. Once you have made your choice, *do not alter it.* Simply raise your hand to indicate that you have put an "*R*" or a "*Y*" in the correct position. When everyone has done that, one or other of these lamps will light up for a short while. Then you make your second prediction, and so on, always trying to anticipate whether the red or the yellow side is going to light up in each particular trial. At first you probably won't have any idea of which to expect, and you will just have to guess, but, after a while, you may find it possible to anticipate the side correctly. Try to get as many correct as possible, but remember, all results are anonymous and there are no prizes. So work inde-

pendently, and make your choice *before* the light comes on. Are you quite clear what to do? Has everyone made their first choice?' (Inspect papers.) 'All right, here is the first trial.' (Flash light for 1 second.) 'Now make your second choice.' An informal chat prior to these instructions reassured the subjects that neither their intelligence nor their personality was being measured, and anonymity was guaranteed.

Method 2. *Verbal material.* The examiner read out a list of 12 adjectives at the rate of one every 20 seconds to small groups of four or five subjects who tried to guess the related adjectives. Of the 50 subjects, 23 came from the same source as the light-guessing experiment, the remaining 27 being undergraduates at University College London, 18 to 22 years of age.

Formal instructions were as follows.

'This is a sort of guessing game. I have a list of adjectives which I am going to read out one at a time. Each of these words provides a clue to the target word which is another adjective related to the clue in some way. When I give you the clue adjective, your job is to try to guess what the target adjective is going to be. So write down the clue adjective as soon as I give it, and beside it write down what you think the target adjective is. You will have 10 clear seconds in which to make your guess and write it down. Then I shall tell you what the target word was, and go on to the next clue. Try to guess what the target is going to be even if you are not very sure, but don't write it down *after* you have been told what it is or the whole purpose of the experiment is lost. Is there anything you are not quite clear about?'

Table 2. *Lists of synonyms and antonyms*

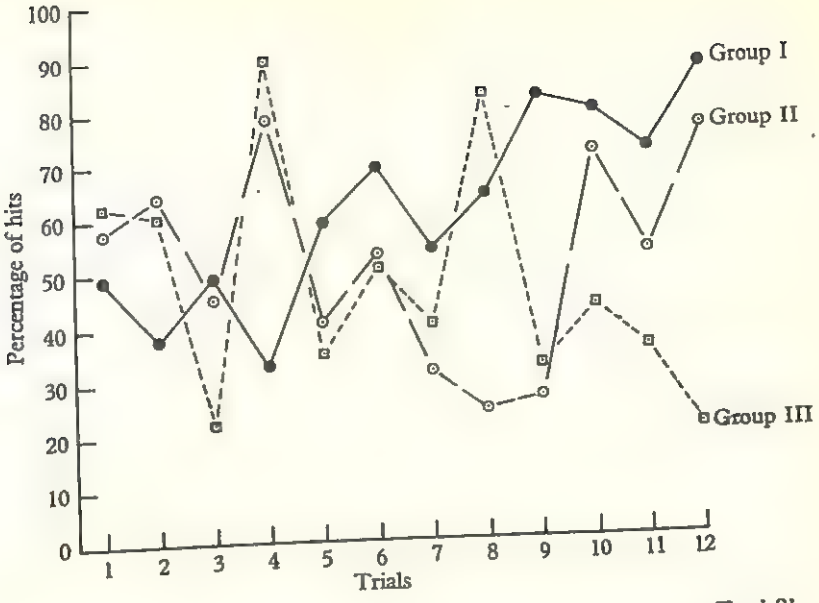
List A Target adjectives		Clue adjectives		List B Target adjectives
Fast	S	Quick	A	Slow
Broad	S	Wide	A	Narrow
Happy	S	Gay	A	Sad
Strong	A	Feeble	S	Weak
Dim	S	Dull	A	Bright
Damp	S	Wet	A	Dry
Brief	S	Short	A	Long
Good	A	Bad	S	Evil
Far	A	Near	S	Close
Low	A	High	S	Tall
Cheap	A	Expensive	S	Dear
Young	A	Old	S	Ancient

S, synonym; A, antonym

Only list structure III, as shown in Table 1, was used, counterbalanced for two groups of 25 subjects. In this case, the binary choice sequence was formed, not with two lights, but with synonyms and antonyms related to the clue adjective. The actual adjectives used are given in Table 2.

III. Results

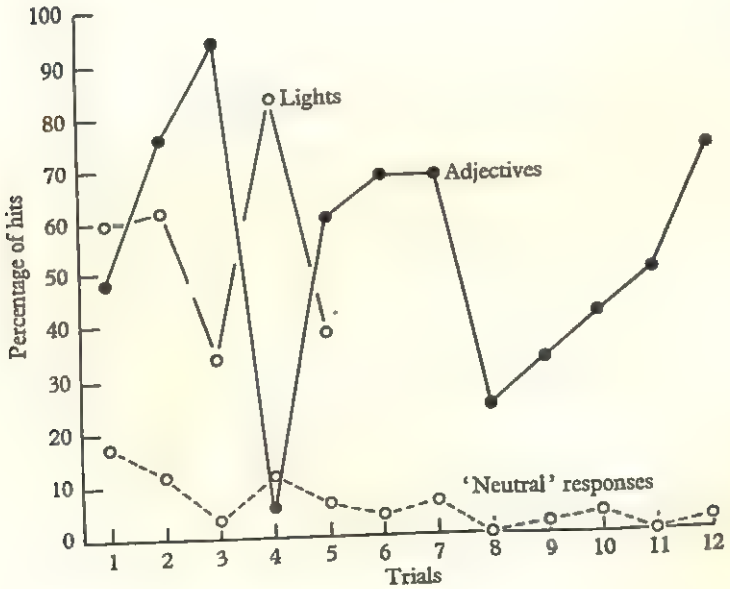
Graphical presentation of the percentage of correct responses for each of the three light-guessing groups in Fig. 1 shows the general trends under these conditions. Group I shows a fairly steady tendency to improve as the simple alternation sequence is picked up. Individual records for this group tend to reveal an abrupt shift from apparently random behaviour to insightful consistency. In the remaining two groups, our interest is mainly centred on the first five responses, which were made under identical conditions for the two groups. The first two guesses both score above chance at about the 60 % level. But having seen the same light come on twice, there is a marked (70 %) mean tendency to swing over to the other response, producing a below chance number of hits. On the next trial, which has been preceded by three instances of the one light appearing, this tendency increases to 80 % and is rewarded. The same tendency persists to a lesser degree (65 %) on the fifth trial. The question is whether, as with Humphreys's (1939) light-guessing data, reinforcement theory could generate the result, or are we committed to some sort of 'expectancy' theory?



Groups													Total S's
I	19	15	19	13	23	27	21	25	32	31	28	34	39
II	24	27	19	33	17	22	13	10	11	30	22	32	42
III	25	24	9	36	14	20	16	33	13	17	14	8	40

Actual numbers of hits

Fig. 1. 'Learning' curves for light-guessing.



Adjs.	24	38	47	3	30	34	34	12	16	20	24	36	50 S's
Lights	49	51	28	69	31	82 S's							

Actual numbers of hits

Fig. 2. Comparison of verbal and light-guessing data.

In Fig. 2 the data for the verbal material are plotted for all twelve trials, and the combined data for groups II and III previously considered are included for the first five trials to facilitate comparison.

Turning to the results for the verbal material we see very different trends emerging from the same pattern of stimuli for the first five trials. From close to a 50% level on the first trial, the tendency to choose the same type of adjective (synonym or antonym) that first appeared increases rapidly to 76% on the second trial, 94% on the third and 92% on the fourth, producing very few hits, as the change in type of adjective occurs on these trials. On the fifth trial, the strong tendency built up on the first three trials drops to the 60% level.

IV. Discussion

The results of group I are of interest for the light shed on the continuity-discontinuity question. No one who considers only the group results would deny the tendency for the scores to move from the random 50% level toward the 100% asymptote. Closer analysis of the individual results reveals a different picture. The issue to be settled is whether the subject shows any gradual improvement on his scores before he hits the correct hypothesis. This is a trickier question than it looks, and dividing the trials into blocks does not answer it. If, for instance we divide the trials into three blocks of four, then both theories predict that the fewest correct will occur in the first block and most in the third. Gestalt theory is covered by the fact that, while a few subjects will pick up the hypothesis early on, more will pick it up during the next block of trials, and so on. Only by examining the results for each individual independently can we circumvent this difficulty. If, before the subject hits the correct hypothesis, he scores above chance, then behaviourists are supported; if the subject scores around chance level they are refuted. How to decide when the subject has hit the correct hypothesis? Ideally, a verbal report should go along with each choice, describing the basis on which the choice was made. These protocols not being available, the next best solution is to adopt some arbitrary criterion in terms of sustained success. Suppose, for instance, we consider that the subject who is correct on the last five trials or more, has solved the problem. The probability of doing this by chance is 1 in 32, or about the 3% level on a one-tailed test. Any subject who does not reach this criterion we consider to have failed to solve the problem, and discard his results, confining our attention to the remaining subjects. For each subject, we consider all trials prior to the final run of hits. If he has succeeded on the last seven trials we count the number of hits and misses for the first five trials. If he gets more right than wrong we score him 'plus', vice versa 'minus', equal numbers 'neutral', and discard the 'neutral' subjects. According to the Gestalt position, the numbers of 'pluses' and 'minuses' should be roughly equal. The behaviourists predict more 'pluses'.

This analysis is set out in Table 3. Nineteen subjects reach criterion and 20 are discarded as having failed to solve the problem. Of these 19, a further 5 are discarded as 'neutral'. None of the remaining 14 scores 'plus'. This is an alarming result since neither party to the dispute makes this prediction (and the sign test shows it to be significant well beyond the 0·001 level of confidence). Something is radically wrong, and, in this case, biased sampling can be convicted as the primary offender. By considering *all* responses prior to the successful one we have ensured that at least one of these, the one immediately preceding the run is a 'minus'. This response must therefore be discarded. The effect of this

Table 3. *Analysis of pre-solution guessing behaviour of nineteen successful subjects*

Total right in final run	No. right before final run	No. wrong before final run	First sign allocation	Second sign allocation	Total wrong in initial run
10	1	1	N	+	0
10	0	2	–	–	2
10	1	1	N	+	0
9	0	3	–	–	3
9	1	2	–	N	1
8	2	2	N	+	1
8	2	2	N	+	1
8	2	2	N	+	1
8	1	3	–	–	2
7	2	3	–	N	0
7	2	3	–	N	1
7	2	3	–	N	0
*6	0	6	–	–	6*
6	2	4	–	–	0
6	2	4	–	–	0
5	1	6	–	–	4
5	3	4	–	N	0
*5	1	6	–	–	5*
5	2	5	–	–	1

* Last discards			Third sign allocation
	+	0	5
	N	5	7
	–	14	7
		19	19

is to reduce all scores in column 3 by one. Re-allocating signs on this basis there are again five 'neutral' discards, but the advantage of 'minus' over 'plus' scores drops to the 10% level of confidence (one-tailed). A difference of this order need occasion no concern, but even this slight trend is attenuated if we consider the following argument. Having been obliged to use an arbitrary criterion of five or more correct in the final run, we take the risk of including in the sample some subjects who have not solved the problem. According to the null hypothesis, these subjects have been responding on a chance basis and therefore would be expected to produce roughly equal numbers of hits and misses. These happen to be distributed so that most of the hits appear at the end of the series and most of the misses at the beginning. Hence, by including these subjects, we bias the sample in the 'minus' direction. It might further be argued that to offset this bias all subjects who get the first five or more in the series wrong should be discarded as non-random by the same token that selects the successful subjects. This procedure eliminates only two of the 19 subjects but reduces the difference between the 'minus' and 'plus' groups to 7:5 and forces the null hypothesis upon us.

So it begins to look as if the gestalt view has the better of the exchange. One fact could save at least part of the behaviourist's case. Suppose we discover from protocols that the subject sometimes verbalizes the correct hypothesis *after* he has been responding correctly for one or more trials. Then, if we base our analysis on the responses previous to the verbal report, we should find a trend in the positive direction, but only *immediately* prior to getting the correct hypothesis. No data are at present available to support this contention, but it looks a tempting interpretation. Unless and until it is confirmed by experiment there is little point in pursuing the theoretical implications of such a two-stage learning process.

When we turn to consider the data for the verbal material, several points need to be dealt with before a useful comparison can be made with the light-guessing data. There is, first, the question of whether we are entitled to regard the verbal task as a binary choice situation. No overt restrictions were placed upon the subjects beyond the hint that the adjectives would be related 'in some way', and the experimenter anticipated some difficulty when it came to marking the data, which would necessitate independent judges to assess whether a response came in the 'same', 'different' or 'neutral' category. In fact, the responses were so unambiguous that no outside judge was called in. As can be seen in Fig. 2 the number of neutral responses, even on the first trial, was only 9 out of 50, and the total in this wastepaper-basket category was 35 out of 600 (or less than 6 %) including complete failure to respond. In this sense, then, the choices were, for practical purposes, binary. The question now arises that, if the subject made a correct response with regard to type of adjective chosen, but chose the wrong specific adjective of this type (e.g. he predicted 'tall' to the clue 'short' when the specific target was 'long') was he reinforced or punished? According to S–R reinforcement theory, if the subject was reinforced, the next time 'short' or something similar appears as a stimulus there is an enhanced probability of 'tall' or something like it being given as a response. Although stimulus and response generalization is an established principle, it is not clear on what basis degrees of similarity are to be computed. The identical elements theory does not apply when dealing with meaningful material. In this case, any adjective is a similar stimulus sufficient to evoke a response that is similar in the sense of bearing a certain *relation* to the stimulus: that of being a synonym or antonym.

The beautiful learning curves that emerge are explicable only provided we accept that it is the *relation* that is being reinforced and learned. If the subject is reinforced for responding 'fast' to 'quick' then there is an increased probability that he will respond 'broad' to 'wide'. Given the same stimuli, but reinforced for saying 'slow' to the first clue, the probable response to the second becomes 'narrow'. Yet there are no grounds for supposing that 'fast' and 'broad' are more similar than 'fast' and 'narrow', or 'slow' and 'narrow' more similar than 'slow' and 'broad'. 'Fast' and 'broad' only become more similar if they are classified as synonyms to the clues. It is the total context that determines the next response, not *specific* preceding stimuli and responses. When synonyms are the stimuli, the data raise other problems for the orthodox S–R reinforcement position, as Marx & Bernstein (1955) have shown.

The data for the two light-guessing groups, taken on their own, are just as inhospitable to an orthodox S–R interpretation. Having been reinforced on the first two trials for saying 'yellow' or punished for saying 'red' there is a reduced probability of saying 'yellow' on the third trial. This paradox is even more apparent with the fourth response. To deal with this result we are obliged to talk in gestalt 'expectancy' terms or else in terms of transfer from previous learning situations, nor is any contradiction involved in talking both ways. To say that the subject expects a binary sequence to show a change after a run of two or three similar outcomes is quite compatible with the statement that the subject is prepared to respond in this way because of earlier training on binary sequences. The gestalt theorist is making a statement of present fact, to which the behaviourist predicates a genetic inference.

We are interested, though, to inquire why it is the subject does not expect a synonym after two or three antonyms or transfer his binary guessing habits to this situation. We

are further interested to know just what the earlier learning situations are that produce these non-random guessing habits. The genetic account is useful only if specific. To the first query we are bound to point out that, although synonyms and antonyms do make a binary choice situation, there is still a certain amount of variety possible within these two classes. In the light-guessing situation, the subject appears to reject the simple repeated response as too easy, whereas some sort of skill is required to find a related adjective. Whether this rejection of the simple solution is innate or acquired is another question. Perhaps the behaviourist would deal with the genetic issue by suggesting that in competitive situations an opponent varies his play as much as possible. This is particularly true of simple two-choice games played by children, such as matching pennies or picking the fist concealing a matchstick. Since these games are played only for a short while, the longer runs characteristic of a random series are inappropriate, and the short run tactic is overlearned.

REFERENCES

BITTERMAN, M. E. (1953). Spence on the problem of patterning. *Psychol. Rev.* **60**, 123–26.

HULL, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychol. Monogr.*, no. 123.

HULL, C. L. (1952). *A Behaviour System.* Yale University Press.

HUMPHREYS, L. G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. Exp. Psychol.* **25**, 294–301.

LAWRENCE, D. H. (1949). Acquired distinctiveness of cues. I. Transfer between discriminations on the basis of familiarity with the stimulus. *J. Exp. Psychol.* **39**, 770–84.

LAWRENCE, D. H. (1950). Acquired distinctiveness of cues. II. Selective association in a constant stimulus situation. *J. Exp. Psychol.* **40**, 175–88.

MARX, M. H. & BERNSTEIN, B. (1955). Generalisation of reinforcement among responses to synonyms. *J. Gen. Psychol.* **52**, 49–64.

MCCLELLAND, D. C. (1943). Studies in verbal discrimination learning. IV. Habit reversal after two degrees of learning. *J. Exp. Psychol.* **33**, 457–70.

PRENTICE, W. C. H. (1949). Continuity in human learning. *J. Exp. Psychol.* **39**, 187–94.

SPENCE, K. W. (1940). Continuous versus non-continuous interpretations of discrimination learning. *Psychol. Rev.* **47**, 271–88.

SPENCE, K. W. (1952). Mathematical formulations of learning phenomena. *Psychol. Rev.* **59**, 152–60.

WALK, R. D. (1952). Effect of discrimination reversal on human discrimination learning. *J. Exp. Psychol.* **44**, 410–19.

(*Manuscript received* 9 *January* 1957)

EXPERIMENTAL DESIGN: A CLOAK FOR INTELLECTUAL STERILITY

By J. G. TAYLOR

Psychology Department, University of Cape Town

Lewin's (1931) warning against applying in psychology the Aristotelian notion that frequency of occurrence is equivalent to lawfulness fell on deaf ears. To-day Aristotelian modes of thought are as prevalent as ever, and are backed up by an elaborate structure of statistical methods and rules for experimental design. By reason of their logical tidiness these methods have the unfortunate effect of reducing (to a level below the threshold for effective action) the discontent that should be engendered by the rotten state of current psychological theory. What psychology needs is not statistical laws but scientific laws, that is, statements, expressed in general terms, that can generate specific predictions when specific values are assigned to the general terms. Such laws are unlikely to be found by statistical investigations, since they are concerned not with populations of systems but with single systems. Thus the law of gravitation is not a catalogue of the properties of solar systems, but it explains how any given solar system hangs together. It is argued that since mind occurs in an organic system its laws are likely to be laws of an ultrastable system (Ashby, 1952), and therefore not discoverable by statistical methods. Two kinds of investigation are required. The first is exemplified by Piaget's studies of the development of mind in individual children, and the second by Kohler's work with distorting spectacles. The first shows how the organism, considered as a multistable system, adapts itself by stages to its environment. The second involves a deliberate disorganization of both behaviour and perception, and enables us to observe simultaneously the processes of behavioural readaptation and the correction of perceptual distortions. It is claimed that the results are compatible with the author's theory of the behavioural basis of perception.

I. Theory and experiment

It has often been said that what is wrong with modern psychology is that there is too much theorizing and too little experiment. Superficially this appears to be true; nevertheless I propose to maintain the opposite thesis, that there is a dearth of serious theorizing and an excess of trivial experimentation. It is true that there are too many theories, but they are not theories in the sense in which the term is understood in other scientific disciplines. A genuine scientific theory is a model that seeks to account for some specific aspect of nature *within the framework of science as a whole.* Behaviour theory satisfies this requirement, but in the field of psychology* proper, the study of mental phenomena, it seems to have been completely neglected. Without the restraint that this demand imposes, psychological theories can be multiplied *ad infinitum.* There are certainly too

* In this paper I need to distinguish between the study of behaviour and the study of conscious phenomena. Originally the term 'psychology' referred to the latter study alone, and it seems highly desirable that it should continue to be so used. When I use the term psychology I wish the reader to understand that I am specifically excluding behavioural science. It is quite possible to write to-day a 'Psychology from the standpoint of a behaviourist', but it would be a book about *mind*, that is, about conscious phenomena. Watson's book with this title was a book not about mind but about behaviour. By using the word psychology in his title he created a lot of unnecessary confusion.

There is similarly a lot of confusion over the use of the word 'mind'. I shall use it to mean the states of consciousness, actual and potential, that characterize an individual. No reference is intended to any ghost in the machine, nor to any joker lurking in the cellarage, whether Freudian, Jungian or anything else, nor to any organizing principle or any other theoretical constructs.

many theories of this kind, but there is a conspicuous lack of theories that can withstand detailed logical scrutiny and involve no conflict with any of the known laws of nature. Indeed the majority of psychological theorists seem to act as if the laws of nature were totally irrelevant to their problems.

The situation is even worse than it was when Murchison published his *Psychologies of* 1925 and *Psychologies of* 1930, for whereas the contributors to those volumes did not hesitate to admit the full extent and gravity of their differences with one another, the contestants of to-day seem to have agreed to forget the basic philosophic and scientific issues that troubled their predecessors, and pretend that those issues are no longer of any importance. Consider, for example, the nativist–empiricist quarrel. At one time the two points of view were totally irreconcilable, and each side understood quite clearly that any concession to their opponents would undermine the logic of their theories. To-day they are all eclectics: there is not a nativist who will deny that there are some aspects of mind that are modified by experience, and there is hardly a single empiricist who will not concede that certain features of mind are innate. The implicit dethronement of logic is apparently a matter of no consequence!

From one point of view this extensive unanimity might be interpreted as meaning that the science of psychology is growing up, in that there is now a substantial measure of agreement among its practitioners. This is a dangerous fallacy. What has actually happened is that, on the one hand, a growing body of evidence has made it impossible to maintain a thoroughgoing nativist theory, and, on the other hand, no really rigorous empiricist theory has yet been published, so that the psychologist with empiricist leanings is constantly being confronted with phenomena that he is unable to deduce from his 'theory'. In these circumstances, neither side is able to deliver a vigorous attack on the other, and this paralysis is sometimes interpreted as meaning that the competing views are less fundamentally opposed than they were earlier thought to be. Certainly, two eclectic points of view that borrow, in somewhat different proportions, from the same mutually incompatible philosophies, are not any more opposed to each other than the hypotheses that the devil is black and that the devil is red. But the philosophies from which both have borrowed are still mutually incompatible. It is impossible for some forms of mind to be innate and others acquired, just as it is impossible for some forms of matter to be composed of electrons and other particles in motion, while other forms are made of particles that are in static equilibrium. If we choose to ignore this fundamental contradiction, dismissing it as a pseudo-problem, we thereby preclude ourselves from asking any vital questions at all.

If we do not allow ourselves to think, there is only one way in which we can create the illusion that we are still honest scientists: we must keep the stream of experimentation flowing continuously. But this creates a new dilemma. The stream of worthwhile experiments has only one genuine source, and that is hypotheses that have some pretensions to being accepted as universal laws of nature. But if we refuse to face the most fundamental problems, we cannot formulate a single hypothesis that can be included in this class. So we must hunt around desperately for a substitute, and the most promising candidate that has appeared so far is experimental design. We can now ask the most trivial questions we like, and provided we set about finding the answer in the approved fashion, using all the new tricks of experimental design, we can be reasonably certain that someone will hail our discoveries as 'important contributions to knowledge'.

My complaint against experimental design is not that it is bad in itself. It is, on the contrary, a most admirable body of doctrine which, in its rigorous self-consistency, stands in strange contrast to the chaos of psychological theory. In fact, it is this contrast that is at once its chief attraction, and the chief source of the damage it does to psychological theory. For if a psychologist can design his experiments in such a manner as to make full use of all the apparatus for testing significance, analysing variance and the like, his demand for logical tidiness is satisfied even if the resulting contribution to science is negligible. The essence of my complaint is that, by partly satisfying the investigator's demand for logical consistency, experimental design protects him from the acute discomfort which the general theoretical chaos should induce in him and thereby inhibits him from lending a hand in the job of cleansing the Augean stables of current psychological theory. The need has been urgent for a long time, but so far the cleansing stream has been little more than a trickle, and for this I put at least a part of the blame on blind faith in the efficacy of experimental design.

This is a strange outcome of procedures that ostensibly were devised for the express purpose of answering theoretical questions, and it is important that we should try to find out what has gone wrong. I suggest that part of the answer is to be found in the origin of some of the statistical techniques that are now so widely used. Consider, for example, the analysis of variance. Originally this was a technique devised to provide answers to certain practical problems in agriculture, such as deciding which of several varieties of a plant might be expected to yield the biggest crop under different sets of conditions. It was not intended as a substitute for basic scientific research, although it could be used as a tool for testing deductions from hypotheses arrived at through basic research. The farmer cannot afford to wait for the slow process of basic research to solve all his problems; he must have answers with a minimum of delay, and carefully devised experiments in which provision is made for the analysis of variance do precisely that.

A similar urgency in the Second World War led to the widespread adoption of the same techniques by applied psychologists. In fact, the urgency here was even greater, for, whereas the agricultural scientist has at his disposal a substantial body of accepted doctrine on soil chemistry and plant physiology, the psychologist has virtually no established theories on which he can build, and therefore must of necessity use methods which do not depend on basic concepts. The successful use of such methods in a wide range of researches on the selection of personnel and the like has added enormously to their prestige and made it almost impossible for a psychologist to ignore them, even if he is engaged in pure research. If he is an 'applied' psychologist there are no other tools he can use, for, ironically enough, there is no body of doctrine which everybody agrees to call psychology that is capable of being applied. The one thing an applied psychologist cannot and dare not apply is *psychology* in any of the forms in which it is at present understood. In the circumstances, the only thing he can apply is statistical method. But this does not justify the extensive use of such methods by psychologists who are engaged in pure research. The kind of answers we have a right to expect from them are not statements of probabilities relating to isolated sets of phenomena, from which we can deduce nothing concerning any other phenomena, but general laws that are applicable throughout the whole field of psychology.

II. SCIENTIFIC LAWS AND STATISTICAL LAWS

The demand for universality, the justice of which no psychologist will question, is commonly understood as meaning that, for a law to qualify to be called universal, it must be shown to produce the *same results* in every person. But this type of generality is precisely what has never been demonstrated in any type of psychological experiment. By this I do not mean merely that results vary quantitatively from person to person. Quantitative variation does not invalidate a general law provided that the variations are themselves accounted for by law. What is much more serious is that there is hardly a single experimental procedure that does not produce results that differ qualitatively. Clearly a law cannot claim to be universal if it predicts one type of result, and experiment shows that there are two or more qualitatively different types of result. Confronted with this uncomfortable fact, there are many psychologists who do not hesitate to conclude that psychology is by its very nature an *inexact science*. This seems to imply a belief not merely that nobody has yet found the laws that will account exactly for all types of results, but that no such laws exist in nature. But this is absurd. Mind constitutes part of the system of nature, and it must be lawful. How then are we to discover its laws? Clearly not by statistical methods, since these do not enable us to calculate exactly the results to be expected in an individual case.

An analogy may reveal what is required. If Newton had had at his disposal not a vast amount of detailed information about a single solar system but a much smaller number of facts about each of a thousand solar systems, collected by a thousand observatories, he might conceivably have developed statistical methods for organizing this material. He might have found correlations between such variables as the number of planets in the system, the average number of satellites per planet, the average distance of the planets from the sun, and the like. He would, by this means, have learned a good deal about solar systems in general, but he could not have calculated the time and place of the next eclipse of the sun, and he could not have arrived at an understanding of the laws of planetary motion. He would have learned a lot about the ways in which solar systems differ from one another, but nothing about the way in which any one of them works. For this it was necessary to know as much as possible about one system. Fortunately Newton had no alternative, and the result of his labours was the construction of a theory that survived until the advent of Einstein's theory of relativity.

If we think of each individual mind as a separate system, then it may be said that the existence of innumerable systems, all differing from each other in detail, and all accessible to observation, constitutes a serious obstacle to the construction of universal laws of mind. The temptation to compare those systems and to think of what they have in common as being lawful, and of deviations from normality as capricious or at least unexplainable in terms of law, is still very strong more than a quarter century after Lewin's (1931) forceful exposure of the futility of this Aristotelian way of thinking. With such a wealth of material at hand it is easy to be misled into thinking that the counting of heads may reveal something that occurs universally and will therefore enable the investigator to write a 'universal' law of nature. But a real universal law has nothing to do with universality in this sense. Newton's first law of motion, for example, states that a moving body continues to move in a straight line unless deflected by some force, but the frequency of straight paths in the celestial sphere is nil. Again, his law of gravi-

tation says nothing about frequency. It simply says that any two bodies of mass M_1 and M_2, separated by a distance d, are attracted to each other by a force that is equal to GM_1M_2/d^2. Now this law of gravitation does not tell us how many planets a solar system must have, what their masses must be, or their distances from the sun, or the shapes of their orbits. What it tells us is that, given certain facts about the sun and planets, we can calculate the behaviour of the system as a whole, and of course the behaviour of any selected unit, including any that may be artificially added to the system. Similarly, the universal laws of mind, when they are discovered, will tell us nothing directly about the actual content of any specific mind. They will be formulated in such a way that, given certain facts about an individual, including facts that are not in themselves mental phenomena, we shall be able to calculate other properties of the system. A law of nature is not a catalogue of the parts that every system must have, or a classification of the parts according to the frequency of their occurrence, but an explanation of how the parts hang together in the system. Thus a law of nature can never be stated in the form of a statistical law, although statistical laws may be derived from it.

Conversely, it will not as a rule be possible to arrive at a law of nature by means of investigations that rely largely on statistical methods. I do not mean by this that these methods are of no value at all. On the contrary, they may, by revealing loose bonds between certain variables, prompt the investigator to search for the connecting links and thus lead to the discovery of the true laws of the system. This, for example, is what Eysenck is aiming at in his investigation of the dimensions of personality. Unfortunately, the majority of psychologists continue their further search by making still more statistical investigations, usually including variables that have hitherto been neglected, and in the end they are as far away from the goal as ever. The reason for this failure is not that there is anything wrong with the design of their experiments but that they have asked the wrong kind of questions. They have neglected the vital question: what makes the system hang together in any one person?

A somewhat crude analogy will help to make the point clear. Suppose that an investigator, knowing nothing about the construction of a motor car, decided to choose as his area of research the behaviour of the speedometer needle, and to this end took a series of readings in each of a hundred different models. Just to make the problem more like a real one we shall suppose that the speedometer dials are not provided with scales, but that the investigator can measure the angular deviation of the needle. Among the variables he might be expected to record are the distances of the accelerator and brake pedals from the floor, the position of the gear lever, the gradient of the road, the direction and velocity of the wind, and, of course, the speedometer reading. He takes a succession of simultaneous readings of all those variables in each car, and then proceeds to examine his data in the hope of solving the riddle of the speedometer needle. At first the material looks completely chaotic. There is no single independent variable that is functionally related to the dependent variable, and he decides to have recourse to statistical analysis. He finds negative correlations between the speedometer reading and (*a*) the distance of the accelerator pedal from the floor, and (*b*) the gradient of the road; and positive correlations with (*c*) the position of the gear lever, and (*d*) the distance of the brake pedal from the floor. He finds significant differences between the speedometer readings when the gear lever is in first, second, third and fourth positions, but the distributions overlap extensively. He now decides to record additional data, such as the weight of the car and its

consumption of petrol, but the riddle remains unsolved. Of course we know the answer. If our investigator will only take independent measurements of the speed of the car he will find that in each system (car) the speedometer reading is a function of speed, but not necessarily the same function in all systems. He will find, moreover, that he can now dispense with statistical methods and can examine each system, considered as a matrix of pointer readings representing the several recorded variables, to determine how it hangs together. He will discover that what he at first took to be evidence of arbitrariness or caprice in his data was actually an artifact arising from the simultaneous examination of pointer readings taken from a hundred different systems. He will find that the same general principles apply to all the systems, but each of them has its own specific set of parameters, with the result that, in Ashby's (1952) terminology, the lines of behaviour of all the systems are different. Continuing to use Ashby's terms, each system is regular and absolute. It is regular because whenever it starts from a given state and a primary operation is applied to it, such as an increase in the gradient of the road or a specific depression of the accelerator pedal, the system will change to another state, and always to the same state. It is absolute because this is true no matter how the given initial state was arrived at.

Whenever we find a *natural* system, as opposed to the man-made system discussed above, that has these properties, we can begin to investigate its laws, and we shall not find it necessary to employ statistical methods or design our experiments in accordance with a plan laid down beforehand by statisticians. There is never any difficulty in devising experimental tests of a good hypothesis. The real difficulty is in finding the right hypothesis.

III. The psychological system

The man-made system discussed above is a dynamic system, and, consequently, its properties cannot be discovered by observing its state at any single moment of time. It must be subjected to the widest possible range of primary operations in all possible combinations, and the state to which it moves after each operation must be recorded. It is only when we have discovered, by this means, that the system is absolute that we can begin to speculate profitably on the laws that hold it together. Similarly, every organic system is dynamic, but it has in addition the property that Ashby has called *ultrastability*. That is, it can, in response to repeated application of certain primary operations, change the internal relations of some of its variables in such a manner that those primary operations no longer tend to throw it out of balance. It equips itself, so to speak, with an automatic gear-change that responds at once to environmental changes.

This property of ultrastability has the most far-reaching consequences for psychology. It implies that the true character of the system can never be discovered by subjecting it to a series of primary operations *after* the process of internal adjustment has been completed. This was long ago appreciated by students of behaviour, and research in this field has been fruitful precisely because experiments have been designed for the purpose of revealing the steps by which the system changes in response to new conditions. In sharp contrast to this, many students of mental phenomena have shown a marked reluctance to ask how the system arrives at the state found in a mature organism. They seem to be content to examine the system as it is now, and to try to understand it by subjecting a large number of individuals to a standard set of primary operations such as tests and

analysing the results statistically. If similar methods had been adopted by students of behaviour we should certainly have understood very little to-day about the laws of learning.

If the psychological system as it exists in a mature subject is the result of a process of adjustment, two different but complementary lines of research are immediately suggested. In the first place, an attempt should be made to discover and tabulate the state of the system at different levels of development. If mind is not an independent system in the Cartesian sense but is an essential aspect of the organic system, then the property of ultrastability leads to the expectation that this type of investigation will reveal an orderly sequence of changes, although not necessarily the same sequence in all individuals, and not necessarily covering the same period of time. The difficulty here is to devise techniques of investigation that will yield reliable results, and this difficulty has been largely overcome by Piaget, whose long series of studies of the development of mind in children reveals just the kind of orderly sequence of changes that might be expected. The second line of research is experimental rather than observational. It involves the manipulation of variables in such a manner as to throw the system out of balance, temporarily destroying its stability. If the system is, as I have supposed, part of an organic system and therefore ultrastable, it should adjust itself to the experimentally altered conditions and correct the induced instability. The first experiment of this kind was performed by Stratton sixty years ago, and he obtained the clearest possible evidence of ultrastability. When he first put on his inverting spectacles the world appeared to swing in front of him in the direction of his head movement but at twice the speed. That is, the positional stability of the visual field, which is a conspicuous feature of normal perception, was destroyed. *But it was almost completely restored in the course of the experiment.* This important result has been conveniently overlooked by critics who have been mainly concerned to prove that Stratton never saw the world the right way up through his lenses. His experiment was not continued long enough to yield unequivocal evidence on this point; but this is not true of the experiments conducted by Erismann and Kohler at Innsbruck, in which stability was so completely restored that the reversed visual field was perceived the right way round.

This clear demonstration of ultrastability in the perceptual field very strongly suggests that mental phenomena are intimately related to the learning process by which the ultrastability discussed by Ashby is achieved. By this I do not mean that perceptual responses are learned in the same way as striate muscle responses. This is the transactionalist view of perception which Pastore (1956) has recently shown to be logically untenable. I prefer the more parsimonious view that perceptual stability is a direct function of the behavioural stability which the organism achieves automatically through learning. That is, mind is a secondary product of the acquisition of an appropriate set of learned responses.

IV. Experimenting on an ultrastable system

There is not, as yet, a very large body of evidence in support of this hypothesis, just because very few experiments have been directly inspired by it. However, in the course of his normal exploratory activities the scientist often brings to light facts whose significance is not immediately obvious, and one of these facts was found in an undeservedly neglected experiment by Howells (1944). The immediate purpose of the experiment was to determine whether it was possible to induce colour-tone synaesthesia artificially, and

the method employed was to present a tonal and chromatic stimulus simultaneously and to require the subject on each occasion to *name* the colour. The colours used were a saturated red and its complementary. We shall refer to them as R and G. After the first 5000 trials, saturated and unsaturated colours were presented in alternation. We shall use the symbols r and g for the unsaturated colours. Two tonal stimuli were used, differing only in pitch. We shall refer to the lower one as 1 and the higher one as 2. In each group of twenty trials the stimulus pairs (1, R) and (2, G) were presented nine times each and the pairs (1, G) and (2, R) once each, the order being randomized. The plan after the first 5000 trials is most easily described by considering groups of forty. For the even-numbered trials in such a group the plan was the same as in the first part of the experiment; for the odd-numbered trials the stimulus pairs (1, r) and (2, g) were presented nine times each and the pairs (1, g) and (2, r) once each. The subjects were paid for their services and were warned that too many mistakes might endanger their employment.

It is easy to see that in this situation the verbal response 'red' became conditioned to stimulus 1, and 'green' to stimulus 2, so that the stimulus pairs (1, G) and (2, R) ultimately tended to evoke the responses 'red' and 'green' respectively. Learning theory enables us to deduce that the number of erroneous responses to (1, g) and (2, r) and even to (1, G) and (2, R) must have increased progressively, since drive was increased as the tally of errors mounted, resulting in larger increments of habit strength each time a correct response was given to (1, R), (1, r), (2, G) or (2, g). The results reported by Howells are in fact entirely explainable in terms of learning theory.

Howells completed his excellent study by performing two experiments which he modestly describes as supplementary but which are in fact among the most important experiments in the history of psychology. I shall describe only one of them. The apparatus for producing the chromatic stimuli was modified so that the subjects themselves could alter the proportions of red and green light reaching the ground glass screen on which the colours appeared. At the beginning of each trial the apparatus was set by the experimenter so as to produce either R or G, and the subject was required to manipulate the control until the light appeared to him to be neutral. If this was done to the accompaniment of stimulus 1 the light which the subject saw as white was invariably an unsaturated green; with stimulus 2 it was an unsaturated red. That is, *the conditioning of colour naming responses to the tonal stimuli had changed the colour perceptions of the subjects.*

It may be objected that this conclusion is not justified in the absence of a control experiment in which no new responses of any kind were conditioned to the tonal stimuli. As it happens the required control was supplied earlier by Kelley (1934) who, oddly enough, described the technique he employed as that of the conditioned response. Actually, what he did was to present certain pairs of tonal and chromatic stimuli a large number of times but without asking his subjects to do anything except attend to them. Since the experimental situation did not provide for the reinforcement of any responses the subjects might make, there was no conditioning of any kind. Consequently, if the changes in colour perception reported by Howells had nothing to do with the conditioning of colour naming responses but were entirely due to the repeated conjunction of one tonal stimulus with one chromatic stimulus, exactly the same kind of results should have been obtained by Kelley. In fact his results were uniformly negative.

These two experiments, together with much corroborative evidence which I shall review elsewhere, are sufficient to justify the formulation of the hypothesis that mind is a

secondary product of the acquisition of an appropriate set of learned responses. If this hypothesis should be confirmed, as I believe it must be, it will obviously have the most far-reaching consequences. It implies, in the first place, that mind is not an independent system with laws of its own, as Descartes believed and as many modern psychologists apparently tend to believe. But, secondly, it does not imply that mind is a useless epiphenomenon, a superfluous decoration that we could easily do without. On the contrary, it has the supremely important property that it provides an immediately available index of the state of adjustment which an individual has achieved. For example, if we find a subject who is not colour blind but who reports as white a red-green mixture that has an excess of red in it, we can infallibly conclude that he has been through some process of conditioning analogous to that which Howells imposed on his subjects. And with equal certainty we can conclude that if a group of subjects are all agreed in calling the same mixture red, they have been through a process of conditioning determined not by some inquisitive experimenter but by the natural properties of the environment in which they grew up.

This idea was vaguely apprehended by Rorschach when he constructed his ink-blot test. He understood that what a subject perceives in an ink-blot is not a function of the stimulus alone but is largely determined by his personality. What he did not understand is that this personality is made up of a large number of specific behaviour tendencies acquired in the course of his life, and that, as yet, we know virtually nothing about the mental concomitants of such tendencies. The vast amount of research that has been inspired by Rorschach does very little to fill this gap, because there is seldom any attempt to trace a specific response or pattern of responses to its behavioural foundations. Rorschach testers, too, have succumbed to the appeal of experimental design.

It now becomes clear that the primary task of psychology is to trace every aspect of mind to its behavioural roots. For this purpose the ready-made recipes of experimental design will be of little value, because the kind of experiment that is required is one in which individuals are subjected over a period of time to some kind of disruptive treatment that may be expected to produce a mental change. But individuals differ widely in the repertory of habits they bring to such an experiment, so that the treatment does not necessarily produce identical results in all. A genuine scientific law must account for the exceptions as well as the rule, and this implies that a successful experiment must involve a more detailed investigation of individual cases than is customary. The experimenter cannot rest content with a demonstration that the treatment produces the expected results in *some* subjects; he cannot have a true understanding of the laws of the system unless he can show that the same laws, operating on a subject who differs from the others in some particular respect, must produce different results.

To illustrate the point I refer, with the permission of Prof. I. Kohler, to an experiment in his laboratory in which the subject wore right–left reversing spectacles for a period of 71 days. Three experiments with similar spectacles had previously been performed, two in Innsbruck and one in Cape Town, and in all three the subjects ultimately reported that the visual field was no longer reversed. This did not happen in the fourth subject, although he wore the spectacles for a longer time than the others. He did, however, achieve the same stability of the perceptual field as the others. This result appears to invalidate the hypothesis that perception is a function of learned behaviour, since the subject was behaviourally well adjusted long before the end of the experiment. He could,

for example, ride a bicycle with ease while wearing the spectacles, although he declined, after one attempt, to master the difficulties they impose on a swimmer. Why then was there no evidence of the expected ultrastability in his perceptual experience? Actually there was evidence of it. The instability of the perceptual field under rotation of the head disappeared as in other subjects; it was only with respect to right–left orientation that he differed from them. The explanation proved to be very simple. In the course of conversation the subject revealed to me that whenever he was asked whether one object was to the right or the left of another object he invariably related the positions of the objects to the blurred edges of his visual field. Since, for a given disposition of objects relative to the median plane of the head, these positions remained invariant throughout the experiment, and since the right and left edges of the field remained invariant relative to the right and left sides of the head, the subject could always report the field to be reversed. As it happened, his previous training had induced in him a sceptical attitude towards the claim that the visual dimensions right–left and up–down could be changed experimentally. Like many others who lean towards nativism he could not swallow this idea, although he had no difficulty in accepting the equally startling claim that the perceptual instability induced by reversing spectacles could be abolished. He was able to satisfy his philosophic inclinations by developing the habit of describing the positions of objects in space in terms of a reference frame defined by his own head instead of an external reference frame. Had he taken his philosophy a little more seriously he would have recognized that the other claim was equally incompatible with it, and would have discovered that the rotations of the visual field induced by movements of the head, if described in relation to the same reference frame, also remain unchanged throughout the experiment.

We may sum up this group of experiments by saying that all four subjects learned to move in the appropriate direction relative to external objects. In addition, the first three learned to give verbal descriptions of the positions of objects in conformity with their overt actions; the fourth subject found a criterion for such descriptions that was logically, though not practically, incompatible with his overt behaviour but compatible with his pre-conceived philosophic ideas. He achieved behavioural stability without endangering his philosophical stability. In the first three all the most important modes of behaviour conditioned to retinal stimulation were transformed in the same direction, and in consequence the visual world was perceived as correctly oriented. In the last subject one sector of behaviour remained unchanged and the corresponding aspect of perception likewise remained unchanged. It will be seen therefore that the behavioural theory of perception is not only not refuted by divergent results of the same treatment but receives added confirmation from them. If the treatment had induced identical behaviours in all the subjects and divergent perceptual results had still been reported, the theory would automatically collapse.

V. CONCLUSION

Psychological theory as it exists to-day suffers from two closely related defects. It is not general enough, and it is not integrated with the main framework of scientific doctrine. As a direct consequence there is no body of doctrine to which every psychologist can subscribe as there is in other scientific disciplines, including the study of behaviour. Behaviourists differ from one another on the interpretation of details, but at least they are agreed in regarding the field of their inquiry as falling within the system of nature,

so that they cannot, for example, postulate mechanisms that are incompatible with the laws of physics, chemistry, physiology and neurology. Psychologists, on the other hand, often give the impression that they feel themselves to be under no obligation to consider such restraints. They seem to think that as long as they say nothing about neurology they are free to spin any kind of theory that has an air of plausibility, without stopping to consider whether it is consistent with neurology or with behaviour theory.

Now it is easy enough to design an experiment to test such a theory, and it is not at all impossible to get statistically significant results that appear to support it. But there are three other conditions that must be satisfied before it can be accepted as a law of nature. (1) It must be logically consistent. That is, it must not generate deductions that are mutually contradictory. In the science of behaviour it is customary to make some effort to satisfy this requirement, but, in psychology, it is hard to find a theory that does not violate it. (2) A valid psychological theory must be consistent with the general body of scientific doctrine. No doubt every psychologist will admit the justice of this demand, but assenting to it is not equivalent to putting it into practice. There are still habits of thought derived from the time when a wide gulf separated *Geistwissenschaft* from *Naturwissenschaft*, and particularly habits that prevent psychologists from seeing mind as an integral part of an organic system having the property of ultrastability. Specifically our second demand implies that there can be no science of mind apart from the science of behaviour. (3) Finally, a valid theory must be stated in a general form. That is, it must be so formulated that when non-contradictory values are assigned to the several variables and parameters of the system it is possible to draw specific deductions that can be tested by observation or experiment.

Every one of these conditions is necessary, and together they are sufficient to guarantee that any theory of mind that satisfies them has some pretensions to being considered a universal law of nature. If we neglect one or more of them in formulating our hypotheses, we can go on producing hundreds of researches that are impeccable from the point of view of experimental design while contributing nothing to the understanding of mind.

References

Ashby, W. R. (1952). *Design for a Brain.* London: Chapman and Hall.

Howells, T. H. (1944). The experimental development of color-tone synaesthesia. *J. Exp. Psychol.* **43**, 87–103.

Kelley, E. L. (1934). An experimental attempt to produce artificial chromaesthesia by the technique of the conditioned response. *J. Exp. Psychol.* **17**, 315–41.

Lewin, K. (1931). The conflict between Aristotelian and Galileian modes of thought in contemporary psychology. *J. Gen. Psychol.* **5**, 141–77.

Pastore, N. (1956). An examination of one aspect of the thesis that perceiving is learned. *Psychol. Rev.* **65**, 309–16.

(*Manuscript received* 25 *February* 1957)

THE EFFECT OF SET ON ACCURACY OF AUDITORY PERCEPTION

By GEORGE A. TALLAND

Massachusetts General Hospital, Harvard Medical School

Two complementary experiments are reported which compare the facilitatory effects of (*a*) sets established in various ways, and (*b*) different conditions of confirmation, on accurate recognition of words presented aurally, and on systematic distortion of words incongruent with the set.

Sets, whether explicitly established by instruction, or implicitly arrived at from cues embedded in the test material, significantly increase the incidence of accurate perception. The presence of incongruent stimuli does not noticeably reduce the facilitatory effect of a set, not even when confirmation is considerably delayed. Wide spacing of confirmatory stimuli is least conducive to the maintenance of a set.

While one set is in operation, and facilitates veridical perception, another set is unlikely to develop, even though conditions for its emergence are furnished by the stimulus material. However, if an initially established and subsequently confirmed set ceases to facilitate correct perception, it will be replaced by another appropriate set; more rapidly if only some of the words presented were congruent with it, more gradually if all were confirmatory.

Explicit sets elicit the highest incidence of distortion in accordance with expectancy, more particularly if confirmation is delayed or intermittent.

In facilitating accurate perception, double sets are less effective than single sets, whether all stimuli are confirmatory or some are incongruent. The differential effect cannot be attributed to the heterogeneous composition of the stimulus material, nor is it manifested in a manner consistent with a model of reciprocal inhibition exerted by the two sets on one another. The results agree with an inhibitory theory of set according to which sets operate by raising the thresholds of perception for all that is incongruent. The relative lowering of thresholds for the congruent would thus be a secondary effect, as distinct from the hypothetical neural model which accounts for this psychological law in terms of short-circuiting of pathways.

I. Introduction

That perception is a selective process has become a commonplace of psychology. Much ingenious experimentation has gone both into demonstrating that the sensory stimuli furnished do not entirely determine what will be perceived, and into isolating those agencies which select and structure, and also distort, that which is perceived. Theories about the selective processes in perception have been formulated in terms of hypothetical neural models (Koffka, 1935; Köhler, 1929) and also of purely psychological laws. Levine, Chein & Murphy (1942), Bruner & Goodman (1947) and Postman, Bruner & McGinnies (1949) demonstrated that we are more likely to perceive that which satisfies a need or value, or more generally a motive, as Postman & Leytham (1951) later suggested. An offshoot or this trend has been the much debated theory of perceptual defence, brought into prominence by McGinnies (1949), which in turn led Howes & Solomon (1950) to lay stress on the more objective selective influences in perception, such as frequency of usage in recognition of words, Bruner & Postman (1949) having previously demonstrated that the unfamiliar and incongruous are less easily perceived. After thus mapping out the various sources of selective sets in perception, the next task was to determine their relative effectiveness. The studies of Postman & Crutchfield (1952) and of Freeman & Engler (1955) have pointed out the complexity of this problem and the need for further

research, suggesting that set—whatever its source in a given situation—and familiarity are the two major selective factors in perception, but leaving open a third place for the structural or formal properties of the stimuli.

II. Purpose of study

(i) *Dimensions of set and of confirmation*

One object of the present study was to contribute to the dimensional analysis of the selective influence of set in perception. In setting himself this task, as well as in examining the implications of the relative effects of multiple versus single sets, the author picked up some hints from Postman & Bruner's (1949) study of multiplicity of set. One dimension of set or expectancy, that of *singleness* versus *multiplicity*, was derived directly from their study. A second dimension in the present research is type of set, distinguished by the manner in which it has been established. An expectancy can be aroused *explicitly* by instruction or warning, e.g. by announcing that pictures of blue flowers will be shown: a set can also be induced *implicitly* by embedding some cue in the situation, e.g. by listing the names of a few Roman emperors (cue), and then following these up with the names of several others (test series). Sets thus established are maintained if they are corroborated, if they help towards accurate or structured perception, or if they satisfy some need. Inappropriate, i.e. unconfirmed, sets also operate selectively until they are relinquished. Needs being as loosely defined as they are, no set can be shown to be motivationally inappropriate. However, the maintenance of a maladaptive, e.g. a paranoid, set is a problem of abnormal behaviour, and need not concern us here. In contrast to appropriate sets, i.e. expectancies confirmed by subsequent events, a set will be regarded as inappropriate when it creates an expectancy of events which will not ensue. More specifically, it will be called a *counter set* when these events could be anticipated by an appropriate set, the emergence of which it prevents or delays. The second dimension, therefore, ranges from *explicit set*, through *implicit set* and no set to *counter set*.

The third dimension is defined by the degree of confirmation. The theoretical extreme values are those of *complete confirmation*, i.e. every event corroborating the expectancy, and complete absence of confirmation. In this study, besides complete confirmation, an intermediate instance will be considered, namely, *partial confirmation* which means that as many instances fail to confirm as do confirm any one set in operation. These instances can be inserted at different rates, so that confirmation is either fairly evenly *distributed*, or widely *spaced* at first, with the bulk presented in one block at the end of the series, or altogether *delayed*, so that all instances of disconfirmation precede the first stimulus congruent with expectancy. Thus we have three variants of partial confirmation.

In accordance with this dimensional model, one object of the study was to determine the relative facilitatory effect of expectancy on accuracy of perception, varying sets along the three dimensions outlined. To some extent, this task amounted to extending, from the visual to the auditory modality, the work done by other experimenters.

(ii) *The mechanism of set*

There are several explanations of how sets or selective factors in perception, whatever they be called, operate. Obviously, some facilitatory process is involved, phenotypically. In a hypothetical model of the nervous system, however, this can be conceived either as

a central mechanism which raises an otherwise subliminal excitation above threshold level, as suggested by Hebb (1949), or as the effective outcome of an inhibitory process which excludes all or some interference, a view put forward by Miller *et al.* (1951). Postman & Bruner (1949), analysing the comparatively lower facilitatory effect of multiple sets, advanced a neurological model of reciprocal inhibition by competing responses. Resorting to a different model, they also suggested the analogy of 'sharpness of tuning' which is maximal with a single set and most blurred when the organism's receptivity is indiscriminate.

Now, there is no reason why a set should tune less sharply in conjunction with another set, if all that happens is the putting into operation of a central neural process. The appropriateness of the argument of reciprocal inhibition must be demonstrated empirically, as must any explanation in terms of interference from irrelevant or incongruent material. There is no *a priori* reason for assuming that the capacity of the organism, either at the receptor or at subsequent stages of perception, is so limited as to accommodate only a single set at maximum efficiency. This explanation, based on work with motor sets, is not necessarily applicable to expectancies established within the boundaries of a concept (which may be as wide as names of famous men, or as limited as South American states). Experiments comparing single and multiple sets have used expectancies defined by a category or concept, a generic term which facilitates the perception of the specific instance. These categories are pre-established, as is, for example, that for blue flowers. The tuning of the set becomes blurred when we add white and purple flowers to the blue, because we have no acquired habit of response to 'blue, white and purple' as against all other colours. By setting into operation the mechanism which lowers thresholds for white, the independent mechanism controlling the threshold for blue need not be affected. On the other hand, if a set operates selectively by excluding or raising the threshold for the irrelevant, then the boundary between congruent and incongruent becomes not only wider after adding white and purple to the blue, but also less clearly defined, and may indeed cease to be a continuous boundary.

A second object of this study was to test the theories of interference and of reciprocal inhibition, put forward to explain the relative inefficiency of multiple sets as selective factors in perception. Further to this, the results of the experiments were analysed in relation to the alternative models of facilitation by lowering thresholds for the congruent, or by raising thresholds for the incongruent.

III. METHOD

(i) *Subjects*

All subjects were adults, members of the medical, psychological and clerical personnel and of several training programmes at the Massachusetts General Hospital, Boston. Altogether 114 subjects participated in this study; 36 in the control series, 54 in Exp. I, and 24 in Exp. II.

(ii) *Procedure*

Conditions for the utilization of mental sets were provided by presenting ambiguous auditory stimuli for recognition. The effect of set was tested by its facilitation of correct vocal reproduction of words presented aurally at the threshold level of intelligibility. All test lists and instructions were recorded and presented on tape by the same voice and at

a uniform level of intensity (except that this level was considerably louder for all instructions). Each subject's individual threshold of intelligibility was determined by means of a pilot list of 24 mixed words which were not drawn from any one category. The subject's task was to repeat these words as he heard them, and the experimenter set the volume and tone controls so that about as many words were reproduced correctly as incorrectly or not at all.

The subject was seated at a distance of about 4 feet from the tape recorder. The experimenter told him that his task would be to repeat aloud words spoken on the tape recorder, and, if necessary, also assured him that this would not be a projective test. The experimenter further informed the subject that all instructions were recorded. The first of these read: 'You will be hearing several lists of words. I shall read them just loud enough for you to hear me. Therefore you may not clearly understand *every* word you hear. Don't let that disturb you. Repeat out loud every word as it *sounds* to you.' This instruction, which always preceded the pilot list, differed in the *control series* to the extent of announcing 'a list' instead of 'several lists'. In the two experiments, each test list was introduced by an appropriate instruction, e.g. 'Now you will hear names of animals', or 'This time you will hear names of birds and names of parts of the body'. When the experiment called for an *implicit set*, the instruction was: 'Now you will hear another list of words.' These instructions were clearly intelligible, no matter how low the volume control was after calibration. The silent interval after each word was long enough for the subject to repeat it aloud three times at the same pace.

The *control series* consisted of one list of 108 words; the two experiments were divided into several test lists. These lists varied in length according to what type of set and confirmation they tested. The basic unit was an array of 12 words, the number of instances presented in confirmation of any one set, and also the number of mixed incongruent words providing for partial confirmation. These mixed words (no two of which, within any array of 12, represented the same conceptual category) will be referred to as *padding*. Arrays confirming sets were drawn from the following four classes: animals, birds, parts of the body, names of food. These will be referred to by their initials: *A*, *B*, *P* and *F*, and the symbol of padding will be *X*.

Test lists for a single set consisted of 12 words drawn from one class, if confirmation was complete, or of 24 words (half of which were padding), if confirmation was partial. Double sets were tested by 24 words, drawn in equal proportion from two classes, providing for complete confirmation; partial confirmation required adding another 12 words of padding. *Implicit set* was induced by a prelist of 6 words drawn from the same class as the test list. No other indication was given that the latter would consist entirely or partly of words representing the same category, or indeed that they would be derived from any one category. *Counter set* was tested immediately after a list had confirmed an *explicit* set, without break or pause other than that between any two words. The test list for counter set was either an array of 12 words drawn from one class, but one different from that for which the subjects was set, or a list of 24 words, half of which were drawn from one class other than expected, the other half mixed. Test lists derived from more than one category were so composed that no more than two words representing the same class, nor more than two padding words should run in sequence. The selection and ordering of words for each test list were based on the findings of the *control series*.

Responses were recorded under four headings: correct; none; incorrect word or

meaningless sound; incorrect word which is congruent with a set in operation. This last class, which Allport (1955) described as 'hypothesis-confirmation by stimulus-deformation', was here called *assimilation*, i.e. assimilation of the stimulus input to expectancy. The number of correct reports within each array of 12 words was scored separately for each subject. Instances of assimilation were summated for each experimental condition across the entire sample of subjects and converted into percentage of total errors, but not averaged for tests. Differences reaching the 1% probability level were considered significant.

IV. THE CONTROL SERIES

The design of the experiment called for 36 words from each of the four conceptual classes and for 60 mixed words. The latter were so selected as to match the classified words they were to be combined with, both in respect of Thorndike & Lorge (1944) frequency count and, as far as possible, in phonetic structure. Since some pilot runs had indicated that word count rating was entirely unreliable for predicting relative auditory intelligibility, this property of the test material had to be determined empirically. For this purpose three control series were run.

Each of these was presented to 12 subjects, and consisted of the same 60 mixed words (X), and 12 words from each class (A, B, P, F). The order of these words was random, except that those representing any one category were evenly spaced, since, in this instance, the establishment of a set was to be avoided. For the same reason only a third of the words from each class was presented to any one subject. By counting the number of correct reports, an empirical rating of intelligibility was assigned to each classified and padding word. On this basis, three matched arrays were compiled within each category so that each contained about the same number of words with highest rating, second highest, etc.; and these arrays were also equated for order. Five arrays of padding were constructed on the same principle.

Altogether, 1231 correct responses were made to the 60 mixed words by 36 subjects, as compared with 898 to the 144 classified words. The former, of course, were presented three times as often, and therefore these results must be viewed in relation to possible totals of 2160 and 1728. In order to render the results comparable with certain findings of the main experiments, they were converted into correct reports out of lists of 12 words, to give means of 6·84 and 6·24. The difference is not significant ($t=1·58$) and, in view of its direction, no allowance need be made for differential intelligibility in testing the effect of set, since the words for which set was established in the two experiments were not inherently more intelligible than those used for padding.

The ranges of correct responses were 5–35 for mixed, 2–11 for classified words, out of possible totals of 36 and 12 respectively. With means of 6·61 (A), 6·11 (B), 5·78 (P), and 6·44 (F), the four classes did not represent different populations of intelligibility, the F ratio being 0·52.

V. EXPERIMENT I

(i) *Design*

The purpose of this experiment was to compare the facilitatory effects of sets on correct perception in terms of the three dimensions of the present study. This required setting up matched conditions for testing: (1) explicit, implicit and counter set with complete confirmation; (2) explicit, implicit and counter set with partial confirmation; (3) single set and double set with complete confirmation; (4) single set and double set with partial confirmation. Complete and partial confirmation, *per se*, were compared only as

related to double set, in this experiment. The conditions of comparison demanded identical test material, i.e. lists drawn from the same class of words. One category was used for each of the first two tests: A for (1) and F for (2); tests (3) and (4) shared the remaining two categories, B and P. These last two tests also had in common tests of double set with complete and with partial confirmation, so that (4) provided data for comparing the interference effects of padding with those of a second set. In both tests (3) and (4) only explicit sets were in operation. Each of the four tests of comparisons entailed three tests of differences between means. Although the three arrays were matched within each category, possible errors arising from differential intelligibility were eliminated by presenting each array, with its associated padding or second array if it had one, under all three conditions compared. For this reason three tapes were recorded so that, for example, the twelve animal names for which an explicit set was established on the first tape, were listed on the second tape so as to test a counter set, and were presented with an implicit set on the third. All test tapes were re-recorded from a master tape. The design of Exp. I is set out in Table 1, with the following code. The letters A, B, P, F, symbolize the four classes of names, X the mixed words or padding; suffixes stand for first, second, third array within each class, and for the five arrays of padding; a line above the code letter (e.g. $\bar{F}$) indicates explicit set; a dot above the letter (e.g. $\dot{A}$) indicates implicit set; and an asterisk (e.g. $\bar{P}{*}A$) indicates the introduction of a counter set.

Table 1. *Design of Experiment I*

Tape I	Tape II	Tape III
$\dot{A}_1$	$\dot{A}_2$	$\dot{A}_3$
$\bar{B}_1X_4{*}\ \bar{F}_1X_1$	$\bar{B}_2X_4{*}\ \bar{F}_2X_2$	$\bar{B}_3X_4{*}\ \bar{F}_3X_3$
$\bar{F}_3X_3$	$\bar{F}_1X_1$	$\bar{F}_2X_2$
$\bar{B}_3\bar{P}_3X_5$	$\bar{B}_1\bar{P}_1X_5$	$\bar{B}_2\bar{P}_2X_5$
$\bar{A}_3$	$\bar{A}_1$	$\bar{A}_2$
$\bar{B}_2\bar{P}_2$	$\bar{B}_3\bar{P}_3$	$\bar{B}_1\bar{P}_1$
$\dot{F}_2X_2$	$\dot{F}_3X_3$	$\dot{F}_1X_1$
$\bar{P}_1{*}\ A_2$	$\bar{P}_2{*}\ A_3$	$\bar{P}_3{*}\ A_1$

Each tape was divided into three sections of approximately equal length, as shown by the broken lines. Since the experiment consisted of eight test lists, the influence of fatigue, of adaptation and, more particularly, of test sophistication had to be considered. It will be recalled that instructions were not completely confirmed in every instance; partial confirmation also implies disconfirmation, and that with massed effect in instances of counter set. It is possible that, after these experiences of disconfirmation, the effect of a set would be different from what it was before the first encounter of padding or counter set. In order to eliminate systematic errors arising from this source, successive subjects began the experiment with different sections; one in the order of recording, the next with § II and finishing with § I, and the third with § III first and § II last. The final design thus became three-factorial, testing differences between (1) the experimental situations compared; (2) the three tapes; and (3) order of presentation. Subjects were assigned in serial order to the nine cells of the design, first going through the three starting positions of Tape I, so that each cell contained a total of six.

(ii) *Results*

The principal results of Exp. I are set out in Table 2. In accordance with the experimental design, four analyses of variance were calculated. The criteria of homogeneity of variance were satisfied in every instance but one, the exception being the test of differences due to different single sets with complete confirmation ($\bar{A}$, $\dot{A}$, $*A$), in which scores under counter set were not representative of the same population as those under explicit and implicit setting. With regard to order of presentation and tape, the variances were homogeneous, and therefore analysis of variance was calculated with only two factors considered. Differences between counter set on the one hand and explicit and implicit set on the other were tested by Wilcoxon's T for unpaired replicates (1949), that between the latter two conditions by Student's t. The last of these tests gave a ratio of 2·46 ($P > 0{\cdot}01$); both T values (2237 and 2515) were below the 1% limit of 2532.

Analysis of variance showed no significant differences due to order of presentation or to tape in this (or in any of the other) test except that between single set with complete confirmation and double set with partial as well as complete confirmation (P). In this analysis, scores on the three tapes differed ($F = 5{\cdot}67$, D.F. $= 2, 135$),† and it also showed the only instance of a significant interaction effect ($F = 4{\cdot}29$, D.F. $= 4, 135$), which was a product of tape and order of presentation. Much significance, however, cannot be attributed to these results because the variances of scores on the three tapes were not homogeneous.

Table 2. *Mean correct response within arrays; total assimilation; and assimilation as proportion of total error, under the various conditions of set and confirmation in Exp. I*

		Mean correct response	Assimilation	Assimilation / Error
Classified words				
Single set, explicit				
Complete confirmation	$\bar{A}$	9·24	26	0·174
	$\bar{P}$	9·54	43	0·323
Single set, explicit				
Partial confirmation	$\bar{B}$	9·37	20	0·140
	$\bar{F}$	8·52	35	0·188
Single set, implicit				
Complete confirmation	$\dot{A}$	8·55	34	0·183
Partial confirmation	$\dot{F}$	8·17	22	0·106
Double set, explicit				
Complete confirmation	$\bar{B}$	8·17	23	0·111
	$\bar{P}$	8·06	39	0·183
Double set, explicit				
Partial confirmation	$\bar{B}$	7·78	29	0·128
	$\bar{P}$	7·46	29	0·118
Counter set				
Complete confirmation	$*A$	7·02	67	0·249
Partial confirmation	$*F$	7·02	22	0·082
Mixed words				
With explicit single set	$X_{\bar{B}}$	3·65	31	0·169
	$X_{\bar{F}}$	3·78	70	0·157
With implicit single set	$X_{\dot{F}}$	4·80	49	0·118
With counter set	X_{*F}	5·55	28	0·082
With double set	X_{BG}	5·48	26	0·071

Differences resulting from type of set, the principal experimental variable, were significant in each of the three-factorial analyses. With D.F. $= 2, 135$, F ratios were: 14·55 between explicit, implicit and counter set with partial confirmation (F); 23·03 between single set and double set with complete and the same with partial confirmation (P); and 6·60 between single set and double set with complete and with partial confirmation (B).

Breaking down further differences within each of the four triads, the following were found significant: explicit or implicit set versus counter set, whether confirmation is complete or partial; single set with complete or partial confirmation as against double set with complete or partial confirmation. In neither instance does the introduction of padding significantly diminish the effect of a double set, nor is the facilitatory influence of an explicit set significantly larger than that of an implicit set. The differences between these two, when confirmation is complete ($\bar{A}$, $\dot{A}$), is significant at the 5% level ($F = 5{\cdot}41$, D.F. $= 1, 50$) but not significant ($F = 0{\cdot}78$) when confirmation is partial ($\bar{F}$, $\dot{F}$).

The means listed in pairs under the first six headings of Table 2 are, in no instance, significantly different one from the other. Under conditions of double set, individual

† D.F. is here expressed by 2 numbers. The first refers to the number of comparisons being made; the second to the total number of instances involved.

subjects tended to recognize about as many items from one category as from the other. With complete confirmation, 14 subjects correctly reproduced just as many names of birds as of parts of the body; 19 reported more of the former, 21 more of the latter, with differential totals of 47 and 41. The mean difference between the two categories for the entire sample was 1·66; with a S.D. of 1·26. With partial confirmation, eight subjects gave the same number of correct responses to birds and to parts of the body, 26 reported more of the former, 20 of the latter, with differential totals of 52 and 41. Mean difference was 1·72 with a S.D. of 1·48.

Testing the effects of the various sets on correct reproduction against the results of the control series, these were found to be significant in every instance. Whether explicit or implicit, single or double, a set makes for a greater number of accurate responses, than does no set; a counter set results in significantly fewer correct responses. Correct report of mixed words listed as padding was significantly less frequent in every instance than correct report of the same words when presented without a set, i.e. in the control series. These scores also differed significantly according to type of set, more particularly between mixed words presented with explicit, implicit and counter set ($X_{\bar{P}}$, X_P, $X_{\dot{P}}$) and between single as compared with double set with complete confirmation ($X_{\bar{P}}$, $X_{\overline{BP}}$), the F ratios being 8·57 (D.F. = 2, 159) and 15·34 (D.F. = 1, 106); the differences in the inhibitory effect of sets faithfully mirror their facilitatory effects, even to the extent of a difference significant only at the 5% level between explicit and implicit single set ($F = 4{\cdot}85$, D.F. = 1, 106).

Each of the five mixed arrays elicited significantly fewer correct responses than its associated classified array of names. Taking the results of the entire sample of subjects, perfect negative correlation obtains between correct reproduction from the classified arrays and their associated mixed words. Summating the two, the means of combined lists range between 6·15 and 6·51, showing no significant differences.

The errors divide about evenly between no response and incorrect response, the latter including instances of assimilation. The 54 subjects refrained from responding to the 144 classified names in 1206 instances and to the 60 mixed words in 978 instances; the corresponding totals for incorrect reports are 1230 and 1006, respectively. The over-all ratio is 494 no responses to 506 incorrect calls, and this varies between 583 to 417 when there is a single explicit set and complete confirmation, and 411 and 589 for mixed words with a double set. There is no consistent relationship between type of set and direction of this ratio which appears to be rather (though by no means predominantly) a function of the words listed.

When a double set was in operation more errors were made by no response under conditions of complete confirmation, and by incorrect calls under partial confirmation. The total incidence of assimilation, however, was slightly larger in the former situation. Table 2 gives details of assimilation from classified arrays; in the case of double set, however, one must also examine the categories to which responses assimilate. Incorrect reports of names of birds are listed as assimilation irrespective of whether they represent names of other birds or of parts of the body, and the incidence of these two is not necessarily equal in frequency. The totals under conditions of complete confirmation were 26 assimilation to birds, 36 to parts of the body; and under partial confirmation, 25 and 33, respectively. In neither instance is the difference significant between assimilation to or from one or the other category.

Comparing incidences of assimilation in proportion of total errors, i.e. in relation to their potential, significant differences were found in several tests. An explicit single set facilitated significantly more assimilation both to and from class *P* than did a double set, whether confirmation of the latter was complete or partial. Under conditions of complete confirmation, 39 words of this class, out of a total of 213 errors, were misreported congruently with one of the sets in operation (*B* or *P*); 36 words were misreported as parts of the body, out of a total of 420. With $\chi^2 = 6{\cdot}82$ and 8·34 (D.F. = 1), both these frequencies are significantly lower ($P < 0{\cdot}01$) than the 43 errors out of 133 faulty responses when parts of the body were presented with an explicit single set.

The totals of 67 and 22 instances of assimilation under counter set combine 'stimulus-deformation' in accordance with the set originally established and confirmed, and with the other set tacitly introduced while the first was still in operation. Between these two, the entry of 67 for complete confirmation ($\bar{P}*A$) divides into 45 food and 22 animal names. For partial confirmation ($\bar{B}X*FX$), over half the instances of assimilation both from classified and mixed words are congruent with the newly introduced set: 13 out of 22 food words, 17 out of 28 mixed words. With a counter set in operation and under conditions of complete confirmation, the median rank of assimilation in accordance with the previous set is 4·8 as against 8·9 for the new set. The corresponding ranks under conditions of partial confirmation are 10·7 and 7·7. This reversal of the positions is notable, even though it may be an artifact of the lists of words which were not identical under the two conditions. It agrees with another finding, namely, that the establishment of a new set contrary to one previously adopted is more gradual and also more definite when confirmation is continuous. Under these conditions, and with a counter set in operation, a division of correct responses into two or three successive phases shows a consistent upward gradient. The increase from 164 to 217 or from 93 to 137 and finally to 151, is significant as tested by χ^2 (7·51, D.F. = 1; and 16·96, D.F. = 2). When both original and counter set are only partially confirmed, this trend is far less consistent.

VI. EXPERIMENT II

(i) *Design*

The purpose of this experiment was threefold: (1) to retest a finding based on Exp. I, i.e. that complete confirmation does not make for significantly more correct responses than partial confirmation; (2) to explore possible differential effects of set as a function of the dispersion of partial confirmation, whether it is *distributed*, as in Exp. I, *delayed*, or *spaced* in relatively wide intervals, with the bulk delayed; (3) to discover whether, in addition to an explicit set, another set does emerge implicitly, if suggested by and appropriate to the padding words.

In Exp. I, different arrays of words were used to test the effect of set with complete as opposed to partial confirmation. In this experiment, both conditions of confirmation were tested with sets for animal names, arrays A_1 and A_3. The duplication entailed by this design provided data to test whether the samples of subjects participating in the two experiments represented the same population in terms of correct recognition of the test words. Food word arrays F_1 and F_3 were used to test the different models of partial confirmation; list B_2P_2 to explore the emergence of a second set. Array F_1 was combined with padding X_1, F_3 with X_3; partial confirmation of set for animal words was *distributed* among items of mixed word list X_2. *Spaced* confirmation (distinguished by suffix '*s*') meant that words congruent with the set occurred in positions 1, 3, 7, 14 and 17 to 24 in a list of 24. *Delayed* confirmation (the code letter of which is suffix '*d*') followed *en bloc* the entire array of 12 padding words. All sets for animal and food names were given explicitly.

Two tapes were recorded. Following the pilot list, the order of test lists was: set for animal names, complete confirmation ($\bar{A}$); set for food names, delayed confirmation ($\bar{F}_dX$); set either for birds or for

parts of the body, a list containing 12 of each (B/P); set for food names, spaced confirmation ($\bar{F}_s X$); and set for animal names, distributed confirmation ($\bar{A}X$). On one tape, A_1 was confirmed completely and A_3 partially; on the other tape, the design was reversed. Similarly, with F_1 and F_3, each being spaced on one and delayed on the other tape. On one tape test list B/P followed set for birds, and on the other a set for parts of the body. Using the symbols, the two tapes were constructed in the following order. Tape I: $\bar{A}_1; \bar{F}_{d1}X_1; \bar{B}/\bar{P}; \bar{F}_{s3}X_3; \bar{A}_3X_2$. Tape II: $\bar{A}_3; \bar{F}_{d3}X_3; \bar{B}/\bar{P}; \bar{F}_{s1}X_1; \bar{A}_1X_2$. Six subjects were tested on each tape in the original, six in the reverse order.

(ii) *Results*

The results of this experiment are shown in Table 3. Tested by mean correct report of animal names, the sample represents the same population from which subjects for Exp. I were drawn ($t=0{\cdot}35$). There are no significant differences due to order of presentation in any of the lists, between the two arrays of animal or the two arrays of food words, between the effects of set for birds or for parts of the body. Furthermore, no significant differences were found between correct reproduction of animal words, whether set for these was completely or partially confirmed ($t=1{\cdot}37$). Differences in correct repetition of classified

Table 3. *Mean correct response within arrays; total assimilation; and assimilation as proportion of total error, under the conditions of Exp. II*

		Mean correct report	Assimilation	Assimilation / Error
Explicit single set				
Classified words				
Complete confirmation	$\bar{A}$	9·25	10	0·182
Partial confirmation	$\bar{A}$	8·62	11	0·136
Delayed confirmation	$\bar{F}_d$	8·21	24	0·263
Spaced confirmation	$\bar{F}_s$	7·54	22	0·206
Partial confirmation	$\bar{B}/\bar{P}$	9·37	8	0·127
Padding				
Mixed words				
Confirmation delayed	X_{Fd}	2·92	51	0·234
Confirmation spaced	X_{Fs}	3·54	34	0·167
Confirmation distributed	X_A	3·71	19	0·095
Classified words				
Confirmation distributed	B/P	3·21	20	0·095

words are not significant, whether confirmation is *delayed*, *spaced* or *distributed* as in $\bar{F}$ of Exp. I, the F ratio being 2·92 (D.F. = 2, 99). The difference between *spaced* and *distributed* confirmation, however, is significant at the 5% level ($t=2{\cdot}37$, $n=77$) and significant beyond the 1% level ($t=4{\cdot}11$) if we compare the first 10 food words presented in either distribution. In no other instance would a similar reduction of the arrays compared yield a significant difference, if complete arrays of 12 did not. Correct report of expected words in a mixed series of birds and parts of the body (B and P) give a mean of 9·37 which is precisely the same as correct reproduction of bird names distributed among padding words. This mean is also significantly higher ($t=8{\cdot}66$ and 9·56, D.F. = 77) than that of either category if presented with a double set. Correct report of classified names, presented with a set for a different category, is certainly no more frequent than of mixed words under similar conditions, nor is it significantly smaller. Assimilation occurs about as commonly under conditions of either experiment, and does not vary with the spacing of confirmation.

VII. DISCUSSION OF RESULTS

Using a different method of presentation, and measuring the effect of expectancy by frequency of success rather than by degree of ambiguity, this experiment has confirmed previous findings concerning the facilitatory influence of sets. Both the empirically most conclusive and theoretically most suggestive results were found in a situation in which more than a single set was appropriate either simultaneously or successively.

Sets emerging in the process of performance facilitate accurate perception as effectively as those expressly introduced so as to structure the situation in advance. Their inhibitory effect on incongruent stimuli is also about equally strong. These implicit sets are possibly not as powerful as explicit sets in distorting the actual input so as to fit into their terms of reference; this is the case, in particular, when some of the stimuli presented quite clearly do not fit, i.e. padding words. The present experiment has not succeeded in exploring the process in the course of which such implicit sets emerge. The cue series, by which these sets were induced, consisted of six words, each of either high or medium rating on basic intelligibility. Therefore, its reduction to five, to four or to even fewer items, does not substantially change the results, even if one considers comparing scores based on only partly identical test material. Further work will be necessary to determine the boundaries along which implicit sets emerge. They certainly do come into operation under conditions like those of this experiment; their development is delayed if previously another set had been operating, and the longer delayed the more consistently that set was confirmed; continued confirmation of such an earlier set, even though only intermittent, utterly impedes the emergence of an implicit set.

Sets are relatively persistent dispositions. The experimental conditions of the present study did not probe their limits, but they have shown that sets are maintained even though confirmation is considerably delayed. There is some evidence that initially loose spacing followed by solid ranks of confirmation is less conducive to the maintenance of expectancy than is evenly distributed or continuous or delayed mass confirmation.

The amount of padding (or incongruent) stimuli used in this study with single and double sets did not effect disconfirmation. It served to demonstrate the interference effect of sets on perception of incongruent material; these effects tend to be proportionate to the facilitatory effect on congruent sensory input, so that the amount of information transmitted, from relevant and irrelevant stimuli combined, is about constant under different conditions of expectancy. This rule, however, only holds in global terms, i.e. for the entire sample of subjects as a whole, not for individual subjects.

Like Postman & Bruner's (1949) tachistoscopic study, this auditory experiment has also concluded that the effect of expectancy is weaker with a double than with a single set. The present study also agrees with theirs in finding that sets, when multiple, exert a reduced effect by assimilation or formulation of congruent 'hypotheses', However, it has not traced a consistent tendency to abstain from responding rather than to report incorrect words under conditions of partial as well as complete confirmation. When incongruent stimuli are included among the confirmatory, the effect of a double set is, naturally, less pronounced than that of a single set, but not significantly less so than that of a double expectancy with complete confirmation. This conclusion corresponds to the finding that the presence of incongruent material does not, of itself, alter the effect of a single set. Since the facilitatory effect of a double set is less than that of a single set

under conditions of partial confirmation, this decrement cannot be attributed to the heterogeneity of the stimuli.

The results fit well into Postman & Bruner's 'hypothesis' theory, and also agree with their view that sets operate by excluding the incongruent as well as by lowering thresholds for the perception of the congruent. In fact, the findings agree best with a model that treats the latter effect as a function of the former which can be conceived as an inhibitory process. That, operationally, sets do exercise an inhibitory effect is clearly demonstrated by the findings that (*a*) odd, i.e. unclassified or mixed, words are less frequently reproduced when incongruent with a set than they are when presented while no set is in operation, and (*b*) once a set has been established and continues to facilitate perception, another set will not emerge spontaneously, even though cues are given far in excess of what would be needed if no other set were in operation, and even though two sets can be effective simultaneously. Of course, failure to report is not the same as failure to perceive. However, our subjects had been instructed to repeat whatever they heard, and the fact that they made rather more errors by incorrect calls than by no response suggests that they observed this rule.

The experimental evidence here presented merely adds to the massive documentation of the inhibitory effect sets exert on that which is incongruent with their terms of reference, that is, of inhibition in an operational sense. Is it this same inhibitory effect which accounts for their relatively lower facilitatory influence, when presented in double or multiple conjunction? Postman & Bruner consider inhibition primarily in this context of rivalry and tentatively go beyond a purely operational definition of the term by hinting at a neurological parallel of responses competing for a common pathway. The same model may, but need not, be extended to account for the effect of a single or multiple set, as against all that for which no set is in operation. The findings of this study are presented in support of this extension of the inhibitory model, but not specifically in support of the analogy of reciprocal inhibition. In accordance with the latter, either of two response patterns to double sets would be expected: (*a*) one of the simultaneous sets becomes prepotent and operates with the same effect as a single set, while the others cease to have a facilitatory influence on perception; (*b*) the two or more concurrent sets each dominate a phase at the end of which they submerge for the duration of another phase or until newly evoked. The former, i.e. the adoption of only one out of two concurrent sets, is characteristic of some psychotic patients we have experimented with. It did not, however, characterize the subjects of this experiment; these subjects accurately reproduced about equal numbers of words from each of the rival categories under a double set, and did this in frequent alternation rather than in longer phases. Lastly, a model of reciprocal inhibition could not account for the increase in correct recognition of irrelevant words under double as against single set, unless it also implied reciprocal inhibition of inhibitory effect.

What reason is there for hypothesizing that sets facilitate perception by virtue of an inhibitory process, by raising thresholds for perceiving the incongruent, rather than by lowering them for the congruent or by summating the peripheral excitation with a central process? If sets operated by this latter mechanism, the introduction of a double set should not result in diminished facilitation, unless we allow for reciprocal inhibition or attribute a weakening of the expectancy effect to interference from incongruent stimulus material. Both these explanations have already been ruled out in the previous discussion. According

to this hypothesis, moreover, an increase of assimilation would be expected from single to double set, since a larger selection of responses—words congruent with each of the multiple sets, as against those drawn from one only—can be elicited without the necessity of confirmation by the sensory input. This is precisely what occurred in the experiment with counter set. No such rise in the incidence of assimilation occurs with a double set; in fact it falls below the level characteristic of a single set. In view of the wider scope offered for assimilation by the smaller total of correct responses under a double set, this result is the more significant. It is difficult to explain in terms of duplication of lower thresholds or of other facilitatory central processes, unless by hypothesizing that the pathways activated in these processes are so limited that they cannot subserve selective behaviour directed toward two or more previously discrete conceptual categories. A brain model of functional equipotentiality or of alternative circuits does not warrant such a structural restriction. Besides, the hypothesis, even if accepted as a working model, would not directly help in explaining why a single set for a category containing $3N$ members could operate at maximum efficiency, but not a double set for two classes each of which consists of only N members. It would appear that only an adequate model of concept formation could help in answering the question raised. Yet, even if one were available, it would beg rather than answer the question, for concept formation itself occurs only when an appropriate set is in operation. The problem still remains: how are the boundaries determined within which the selective effects of sets operate?

The model presented as being in accordance with the empirical findings of this study (i.e. a model of general inhibition of perceiving all that is irrelevant) is *prima facie* unattractive, for it is uneconomical and contrary to well-established ways of thinking. In fact, it is not inconsistent with previous psychological theory. It makes sense in the analysis of behaviour in certain neuro-psychiatric diseases, and it accords with some recent developments in neuro-physiology, exemplified in the auditory modality by the findings of Hernandez-Péon & Scherrer (1955). Very probably, the hypothesis here advanced will be tenable only in combination with some complementary theory of direct facilitation. It is presented in the most tentative manner; more experimental work along lines similar to those of this study, and others converging from different directions, is needed before this point of view can be stated with greater confidence, or rejected.

REFERENCES

ALLPORT, F. H. (1955). *Theories of Perception and the Concept of Structure.* New York: Wiley.

BRUNER, J. S. & GOODMAN, C. D. (1947). Value and need as organizing factors in perception. *J. Abnorm. (Soc.) Psychol.* **42**, 33–44.

BRUNER, J. S. & POSTMAN, L. (1949). On the perception of incongruity: a paradigm. *J. Personality*, **18**, 206–23.

FREEMAN, J. T. & ENGLER, J. (1955). Perceptual recognition thresholds as a function of multiple and single sets and frequency of usage of the stimulus material. *Percept. Mot. Skills.* **5**, 149–54.

HEBB, D. O. (1949). *The Organization of Behavior.* New York: Wiley.

HERNANDEZ-PÉON, R. & SCHERRER, H. (1955). Habituation to acoustic stimuli in cochlear nucleus. *Fed. Proc.* **14**, 71.

HOWES, D. H. & SOLOMON, R. E. (1951). Visual duration threshold as a function of word probability. *J. Exp. Psychol.* **41**, 401–10.

KOFFKA, K. (1935). *Principles of Gestalt Psychology.* London: Kegan Paul.

KÖHLER, W. (1929). *Gestalt Psychology.* New York: Liveright.

LEVINE, R. I., CHEIN, I. & MURPHY, G. (1942). The relation of the intensity of a need to the amount of perceptual distortion. *J. Psychol.* **13**, 283–93.

MCGINNIES, E. (1949). Emotionality and perceptual defense. *Psychol. Rev.* **56**, 244–51.

MILLER, G. A., HEISE, G. A. & LICHTEN, W. (1951). The intelligibility of speech as a function of the context of the test materials, *J. Exp. Psychol.* **41**, 322–35.

POSTMAN, L. & BRUNER, J. S. (1949). Multiplicity of set as a determinant of perceptual behavior. *J. Exp. Psychol.* **39**, 369–77.

POSTMAN, L., BRUNER, J. S. & MCGINNIES, E. (1949). Personal values as selective factors in perception, *J. Abnorm. (Soc.) Psychol.* **42**, 142–54.

POSTMAN, L. & CRUTCHFIELD, R. S. (1952). The interaction of need, set, and stimulus-structure in a cognitive task. *Amer. J. Psychol.* **65**, 196–217.

POSTMAN, L. & LEYTHAM, G. (1951). Perceptual Selectivity and ambivalence of stimuli. *J. Personality*, **19**, 390–405.

THORNDIKE, E. L. & LORGE, I. (1944). *The Teacher's Word Book of* 30,000 *Words*. New York: Columbia University Press.

WILCOXON, F. (1949). *Some Rapid Approximate Statistical Procedures*. New York: Amer. Cyanamid Co.

(*Manuscript received* 27 *February* 1957)

TACHISTOSCOPIC EXPOSURE TIME AND SPATIAL PROXIMITY IN THE ORGANIZATION OF VISUAL PERCEPTION

By J. W. KASWAN

Psychology Department, Yale University

This study investigated the temporal sequence of spatial organization, using varying proximity of geometric designs (e.g. (h) .. .. .. and (d) .. .. ..) as the criteria for organization. In addition, the study explored the possible impairment of such space-time relations in schizophrenics. Perception of groupings evolved in a predictable sequence of differentiation, correct groupings (e.g. pairs of dots) emerging more quickly the better the spatial differentiation. The results were similar for twenty-four medical-surgical and twenty-four schizophrenic hospital patients, except that the latter required more time to see correct groupings in most of the designs.

I. Introduction

It is generally recognized that visual perception occurs in time, and the literature of recent years abounds in studies on the relation of temporal factors to intensity of stimulation (Hunter & Sigler, 1940), the span of 'discrimination' (Gilliland, Hofeld & Eckstrand, 1946), and other phenomena. The present study concerns two related problems. One is the relation of exposure time to the phenomenology of spatial organization, as characterized by simple geometric patterns varying in the proximity of grouped units. The other is the possible impairment of the time sequence of perceptual organization in schizophrenics.

As regards the first problem, it is generally agreed that some of the grouping factors emphasized by Gestalt psychologists play an important role in perception. Some of the designs originally used by Wertheimer (1950, p. 88) to demonstrate the role of grouping in perception were adapted by L. Bender for inclusion in a test designed to measure perceptual-motor integration. She found that when these designs were presented to normal subjects at fast tachistoscopic exposures the resulting percepts were similar to those observed in small children and some schizophrenics, for whom exposure time had not been reduced. Her work led her to suggest that a time factor may be involved in the organization of perception and that this organization may follow some orderly sequence of differentiation in which a dextrad horizontal ordering of events is an early stage (Bender, 1938, pp. 51–52). Since visual perception usually involves complex groupings of details, determination of time parameters and their relation to spatial characteristics is important not only for a better understanding of perception, but perhaps also for the elucidation of some aspects of behaviour in psychopathology. Thus, there is evidence that schizophrenics show much longer reaction times than normals in a variety of tasks and a basic disturbance of adaptive capacity has been postulated to account for this deficit (King, 1954, p. 155). Since perceptual activity is the first link in the organism's relation to the environment, an increase in the time required for perceptual organization may well be part of the reason for the general sluggishness of reactive behaviour characteristic of such patients.

With spatial proximity as the criterion for perceptual organization, it was assumed that perception of groupings in time would follow a process of differentiation. Specifically, it

was proposed that: (*a*) the clearer the grouping, the less time should be required to perceive it; (*b*) schizophrenics should require a longer time to perceive correct groupings; and, (*c*) the process of differentiation should generally proceed in the following sequence: indistinct blur (e.g. bunch of dots)—dextrad horizontal line of objects—incorrect or unclear groupings—correct groupings.

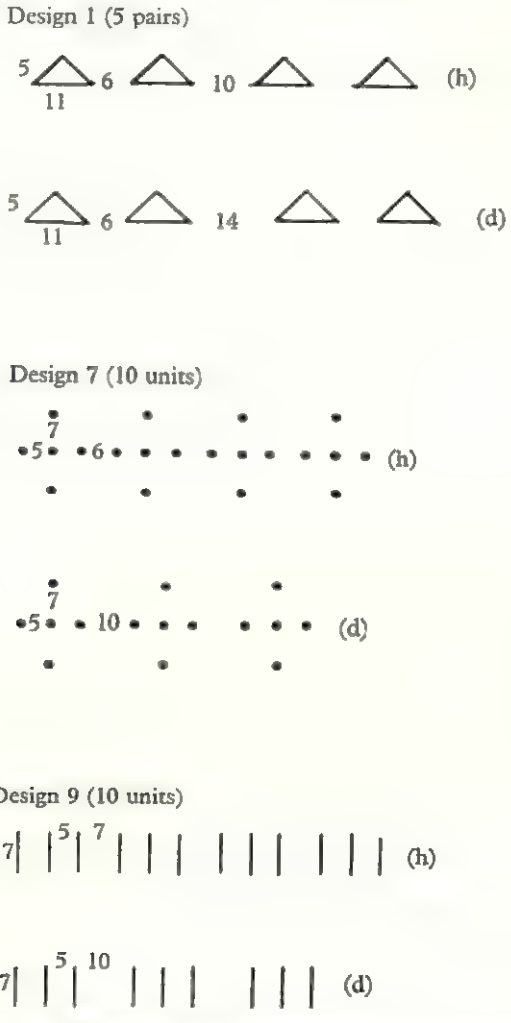


Fig. 1. Examples of designs. Numbers refer to measurements in millimetres of actual size.

II. Methods and procedure

(i) *Materials.* There were a total of twenty-one different designs, some of which are shown in Fig. 1.* All designs were drawn through the middle of 9×12 in. white cards, across the field of vision in the tachistoscope, while a 4 in. margin at the right and a $1\frac{1}{2}$ in. margin at the left fitted into the frame of the apparatus. There was a practice series of nine designs and the main experimental series of twelve basic designs. Ten of the twelve designs consisted of two variants differing in the proximity of grouped units, making a total of twenty-two experimental designs. Where the groupings were closer together, the designs

* To save printing costs samples of all the designs used and the predicted A, B and C responses for each have been deposited with the Technical and Documentary Unit, Department of Scientific and Industrial Research. The materials may also be obtained directly from the author.

were deemed more homogeneous (h); where further apart, they were better differentiated (d). Two were control designs containing simple line drawings without intended grouping and were included, in part, to check on possible guessing tendencies. Of the nine practice designs four were (h)-type drawings, three were (d)-type drawings, and two were control designs containing no grouping. The designs were selected on the basis of the results of a pilot experiment with about thirty drawings and were those the subjects could most clearly describe and reproduce.

To determine whether the intended groupings were easily and clearly visible in both (h) and (d) versions when sufficient time was allowed for their perception, the experimental designs were presented to two groups of 15 adult subjects who were asked individually to describe each drawing within 5 sec. and then rate them 'very clear', 'fairly clearly', or 'not clear'. Each group was given a random arrangement of five (h) and five (d) designs. The results indicate that 97·2% perceived (h) designs as 'very clear' and 95·5% saw (d) designs as 'very clear', thus demonstrating the unambiguous nature of the designs.

(ii) *Subjects.* There were forty-eight male subjects, all Veterans Administration hospital patients, divided into two groups of twenty-four each, classified as Medical and Surgical (MS) and Schizophrenics (Sch). Selection was by chart diagnosis and no subjects with known brain disturbance or visual difficulties were selected. The schizophrenic subjects were all described as 'chronic' and the ward psychiatrist believed none to be in a state of complete remission. Subjects who had undergone electric shock therapy within 3 months prior to the experiment were excluded. The two groups did not differ significantly in age, education, or intelligence as measured by the Wechsler–Bellevue Block Design and Vocabulary subtests.

(iii) *Apparatus.* A Gerbrands mirror tachistoscope was used on which exposure time could be set from 0·01 to 1 sec., at 0·01 sec. intervals. (This instrument provides a constant level of illumination.) For periods exceeding one second a stop-watch was used. A memory drum was adapted as a recording device on which subjects could draw what they saw without seeing their drawing. The experimenter turned a knob which moved the position of the continuous strip after each drawing.

All subjects' eyesight was tested with a standard Snellen chart and only those with monocular vision of 15/30 or better continued in the experiment.

(iv) *Procedure.* After the test of visual acuity, the practice designs were presented in the tachistoscope with the following instructions. 'In a moment I will show you some drawings of lines or dots in different arrangements. As soon as each drawing appears tell me what it looks like as carefully as you can and at the same time draw it right here. As you won't see what you are drawing, just try to draw so that I can get a good idea of what you saw. Any questions? Let's try it.' Subjects who hesitated were urged to give their first impression; any response delayed more than 5 sec. after presentation was scored as failed. The nine cards of the practice series were all presented first at 0·5 sec. exposure and then again, in the same sequence, for 5 sec. The purpose of the practice series was not only to acquaint subjects with the task to be performed, but also to induce a set to see groupings in the designs. To increase the probability that all subjects had a substantial set to see groupings, only those remained in the experiment who could accurately perceive the correct grouping in the last three (h)-type designs of the practice series.

The experimental series was introduced after the practice series without further instructions. As can be seen from the description of the materials, every design had two variants, differing in the distance between grouped units. Each variant was presented separately, so as to get relatively independent results for the two, and it was therefore necessary to divide the twenty-two designs into two groups, series A and B, each containing 5 (h), 5 (d), and one control design. Designs 1h–5h, 6d–10d and 11 were in series A, the remainder in series B. All subjects were given both series at a 3- to 7-day interval between presentations. Subjects were told nothing about the second presentation at the time of the first. Half the subjects in each group were given series A first, the other half were given series B first.

Six exposure times were selected for each design on the basis of the responses of normal subjects in the pilot experiment, divided into six percentage groups. This was done to minimize the possible effect of rigidity on the part of schizophrenic subjects which had been demonstrated in the pilot experiment. Exposure times varied from design to design but were the same for (h) and (d) variants of the same design. All cards were first presented to each subject at the fastest exposure time (0·01 sec.). Then they were rearranged in a different order and presented at the next, longer, exposure time, and so on, to the sixth exposure time. Exposure times, in seconds, for the designs shown in Fig. 1 were as follows: design 1, 0·01, 0·04, 0·10, 0·25, 0·75, 5 and 10; for design 7, 0·01, 0·06, 0·15, 0·25, 0·75, 5 and 25; for design 9, 0·01, 0·06, 0·10, 0·50, 0·90, 5 and 20. The order of exposures followed an 11 × 11 Fisher Latin square. Serial position was also different in the two series shown to each subject in order to avoid exposure of

Bureau Edn. Psy. Research
... COLLEGE
Dated ...

(d) and (h) variants in the same sequence as the previously shown (h) and (d) variants, but there may have been occasional coincidences. It was hoped that this procedure would minimize any possible learning due to previous experience with the designs.

(v) *Scoring.* General criteria, applicable to all designs, were as follows. 'A' scores: (*a*) no grouping at all was perceived; (*b*) the response was entirely unrelated to the stimulus. 'B' scores: (*a*) perception of some vague groupings; (*b*) groupings that were partly incorrect. 'C' scores: accurate description of the shape and contents of the design. These were the only correct responses and, unless otherwise noted, all statistics are based on the first of three successive 'C' responses to every design. In addition, there were specific scoring criteria for each design. For instance, design 9 (see Fig. 1) has the following criteria: 'A'—line of vertical lines, 'B'—groups of two or four lines; and 'C'—groups of three vertical lines. Responses were recorded verbatim and scored after each individual session.

III. Results

In the practice series there was no consistent difference between the groups for perception of the correct grouping on the different designs at the 0·5 sec. exposure. At the 5 sec. exposure, all subjects who continued in the experiment gave correct responses, with one exception in each group for one of the early practice designs. Four schizophrenics were eliminated from the experiment because they could not meet the minimum requirement, which was to give an adequate verbal description of the correct grouping of the last three (h) designs in the practice series, and it seems probable that all subjects had a set to see distinct groupings in the main experiment.

Table 1 summarizes the significance of the mean difference between the ordinal position of exposure at which correct perception of (h) and (d) designs occurred. It was assumed that it would take longer to perceive (h) designs (higher ordinal position) and the results clearly support this hypothesis, as the mean difference from zero is significant in the

Table 1. *The significance of the mean difference between the ordinal position of the correct perception of groupings in normals and schizophrenics*

	Normals		Schizophrenics	
Designs	Mean diff.	*t*	Mean diff.	*t*
1h–1d	2·33	6·51***	2·08	7·25***
2h–2d	1·33	5·81***	1·46	3·47**
3h–3d	2·01	5·30***	2·17	5·61***
4h–4d	1·62	3·63**	2·75	5·86***
5h–5d	2·62	6·28***	3·08	7·32***
6h–6d	1·08	2·91**	1·67	4·41***
7h–7d	1·50	3·77***	1·92	4·27***
8h–8d	0·33	0·82	1·67	3·74***
9h–9d	2·04	5·70***	2·50	6·13***
10h–10d	0·67	1·77*	0·54	1·64

* $P > 0{\cdot}05$. ** $P > 0{\cdot}01$. *** $P > 0{\cdot}001$.

expected direction in eighteen out of twenty comparisons. Fig. 2 shows the occurrence of 'C' responses in actual time units for the designs given in Fig. 1. While the gradient of the curves varies with the design, the difference between (h) and (d) variants is consistently maintained, most curves approximating the ogive form characteristic of psychophysical functions. Table 2 shows that for thirteen out of twenty Kolmogorov–Smirnov comparisons (Goodman, 1954) schizophrenics required a significantly longer exposure time to see the correct groupings. All differences were in the expected direction. Almost

all 'A' and 'B' responses corresponded to the *a priori* criteria established for each design. Only 3·24% of these 'primitive' structure responses seemed idiosyncratic and most of these were given by three schizophrenic patients.

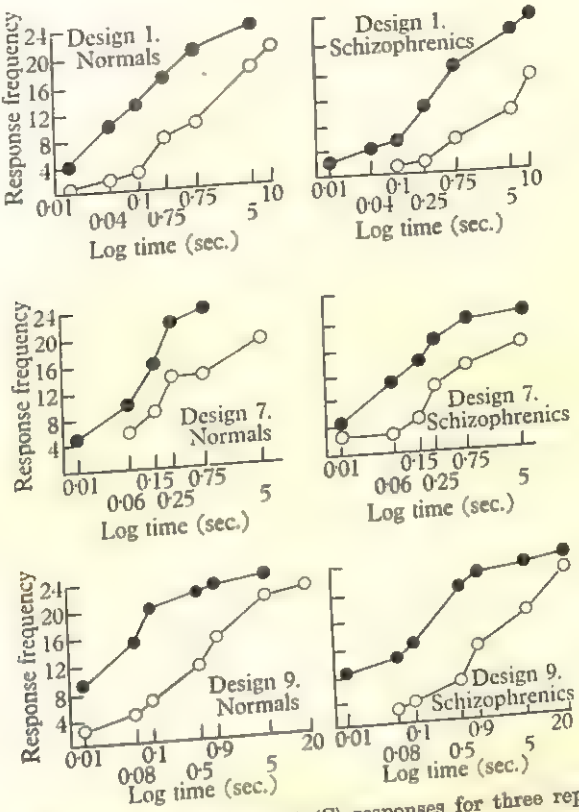


Fig. 2. Cumulative frequency distribution of correct (C) responses for three representative designs. *Key:* blacked circles represent responses to (d) variants, open circles refer to (h) variants of each design.

Table 2. *Kolmogorov–Smirnov test for the significance of the difference between ordinal position of correct perception of groupings in schizophrenics and normals*

Design	$D \times N$	P	Design	$D \times N$	P
1h	9·12	0·03	6h	9·02	0·03
1d	8·16	0·08	6d	5·04	0·30
2h	10·08	0·02	7h	3·84	0·40
2d	9·12	0·03	7d	5·04	0·30
3h	11·04	0·01	8h	10·10	0·02
3d	12·24	0·01	8d	4·01	0·40
4h	9·12	0·03	9h	5·04	0·30
4d	9·01	0·03	9d	8·16	0·08
5h	13·92	0·001	10h	12·96	0·01
5d	9·95	0·03	10d	10·08	0·02

Initial responses were usually 'A' responses, followed by 'B' and 'C' responses. This is suggested by Fig. 3 which indicates a steady decrease in the amount of 'A' responses from the first to the last exposure, with the reverse trend in the case of 'C' responses.

As expected, 'B' responses tend to be perceived later than 'A' responses. All the differences are statistically significant. Comparison of the 'A' and 'B' responses yield a χ^2 of 20·51 for the medical-surgical subjects and of 12·90 for the schizophrenics, significant

at beyond the 0·01 and the 0·05 level respectively. The difference of the trend in 'C' responses compared with 'A' and 'B' responses is too obvious to require computation.

Since the sequence of exposures was from fast to slow, it seems reasonable to suppose that the development of groupings as here described follows the predicted process of differentiation from homogeneity. Since exposures were continued until there were three successive 'C' responses, 'A' or 'B' responses occasionally intervened between the first and successive 'C' responses. The results show that 3·3% of medical-surgical and 11·8% of schizophrenic responses were of this type. The relatively larger number of such responses by schizophrenics support frequent findings that these subjects tend to be more erratic than non-psychiatric subjects.

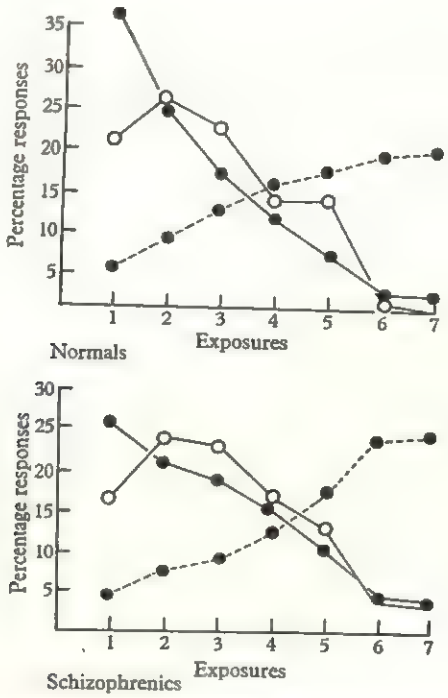


Fig. 3. Percentage of A, B and C responses at each exposure. *Key*: blacked circles with continuous line responses represent A responses. Open circles refer to B responses and blacked circles with dashes refer to C responses.

The two control designs, one each in series A and B, which contained no grouping, were correctly seen at the first exposure by 94% of subjects, indicating that the time factor in the present investigation refers to groupings and not to the perception of figures as such.

Correct perception of corresponding (h) and (d) variants is highly and significantly correlated. The Pearson *r* for medical-surgical subjects is 0·804 and for schizophrenics 0·982. This seems to indicate that the designs used were fairly constant stimulus patterns, changes in proximity of (d) variants affecting performance only in relation to (h) variants.

IV. DISCUSSION

The first problem of this study concerns the relation of exposure time to the phenomenology of spatial organization. The results clearly suggest that the perception of spatial groupings is predictably different at various time exposures. This is demonstrated in two ways. On the one hand it took significantly longer to identify correct groupings in (h) than (d) designs, supporting our expectation that decreased proximity between groupings, favouring horizontal line arrangements, would impede perception of groupings under reduced time exposure relative to (d) variants. On the other hand, it was shown that at short exposure times subjects perceived 'a bunch of lines', or the dextrad horizontal ordering of elements, such as 'bunch of vertical lines', ('A' responses) where vertical lines in groups of three (design 9) would have been the correct percept. At longer exposures in the tachistoscope some grouping was perceived, such as 'dots in groups of threes with some scattered on top and on the bottom' (design 7), and finally, at even longer exposures, most subjects perceived the correct grouping. Thus the results indicate not only that perception of groupings develops in time, but, what is more important, they also provide some clues as to *how* this development occurs. These findings gain added validity from the fact that the differences between (h) and (d) are consistent for two groups of subjects who differ in the total amount of time each required for the perception of correct groupings.

The second problem of this study concerns the schizophrenics' impairment of space-time relations in perception. The longer times required by these subjects to see the correct groupings support the assumption that their long response latencies extend to this very basic level of the organism's translation of the environment. Acquisition of knowledge about the environment within a limited range of inter-individual constancy—perceiving what most other people see—must be at least partly what clinicians call 'sense of reality'.

The time relations explored in this study should be clearly distinguished from those often described by psychoanalysts when they discuss time functions in psychopathology. They speak of the schizophrenic's inability to delay expression of needs (Bonaparte, 1940). The present study deals with time disturbance in the *opposite* direction—the slowness with which schizophrenics perceive groupings. The former involves primarily a defect in the inhibition of need expression, the latter, sluggishness in the pacing of reactions to the environment. While this illustrates that different time parameters may play very different roles in behaviour, only further experimental work—now in progress—can hope to clarify their relationship, as well as their connexion with other cognitive functions, such as learning.

This report is based on a dissertation submitted to the Graduate School of Washington University, St Louis, Missouri, in partial fulfilment of the requirements for the Ph.D. The writer is greatly indebted to Dr Saul Rosenzweig, under whose guidance this project was undertaken. Grateful acknowledgement is also made to Drs Marion E. Bunch, R. M. Collier and Frederick H. Kanfer for their valuable suggestions and criticisms. The cooperation of Veterans Administration hospitals and clinics in St Louis and Jefferson Barracks, Missouri, and Little Rock, Arkansas, is gratefully acknowledged.

References

Bender, L. (1938). *A Visual Motor Gestalt Test and its Clinical Use.* New York: Amer. Orthopsychiat. Association.

Bonaparte, M. (1940). Time and the unconscious. *Int. J. Psycho-Anal.* **21**, 427–68.

Gilliland, A. R., Hofeld, J. & Eckstrand, G. (1946). Studies in time perception. *Psychol. Bull.* **43**, 162–76.

Goodman, L. A. (1954). Kolmogorov–Smirnov tests for psychological research. *Psychol. Bull.* **51**, 160–8.

Hunter, W. S. & Sigler, N. (1940). The span of visual discrimination as a function of time and intensity of stimulation. *J. Exp. Psychol.* **26**, 160–79.

King, N. E. (1954). *Psychomotor Aspects of Mental Disease.* Cambridge, Mass.

Wertheimer, M. (1950). Laws of organization in perceptual forms. In *Sourcebook of Gestalt Psychology*, E. D. Ellis (ed.). New York.

(*Manuscript received* 28 *March* 1957)

PSYCHOMOTOR CONTROL IN STAMMERING AND NORMAL SUBJECTS: AN EXPERIMENTAL STUDY

By GEORGE SETH

Psychology Department, The Queen's University, Belfast

Voluntary control of the regularity of a motor performance (finger-tapping) which does not involve speech or the speech mechanism was investigated in fifteen stammering and fifteen non-stammering subjects. The results indicate: (i) that in right-hand finger tapping the stammerers are markedly inferior to normal controls in their ability to maintain voluntarily the rhythm or regularity of the performance; (ii) that this group of ostensibly right-handed stammerers showed a marked inferiority in right-handed, as against left-handed, performance; (iii) that the disorganization of psychomotor performance is not confined to the speech-function, and may appear in situations where communication by speech is not involved, and in which the influence of personal-social factors is reduced to a minimum.

I. Introduction

It has long been argued by neurological and neuropsychological investigators (Travis, 1931) that the emotional and situational factors that are so obviously influential in stammering speech are secondary and that 'in the great majority of cases the basic cause of stuttering (stammering) is to be sought in some significant variation in the functional integration of the central nervous system'. This view of the stammering disorder, in terms of a physiological rather than a psychological mechanism, derives from a theory of cerebral dominance that has been heavily criticized (Burt, 1937), but it has continued to influence research and treatment up to the present time. In particular, it implies, first, that disturbance in the control of the speech-mechanism should be apparent not only in the speech-situation, and secondly, that in a considerable number of cases of stammering it should be possible to demonstrate inadequate development of the normal preferential sidedness (handedness); or interference, in the 'natural' left-hander, with its establishment. The experimental studies of Blackburn (1931) and Seth (1934) provided evidence for the first of these two hypotheses. These reports showed quite unequivocally that, even in situations in which speech is not required and in which the emotional temperature has been kept as low as possible, control of the mechanism of speech and especially of that of respiration is significantly more erratic in stammering subjects than in normal-speaking controls. Blackburn's results indicated that voluntary control of movements of the diaphragm, lips, jaws and tongue were all defective in the stammerer, and further evidence (Seth, 1934), while showing no statistically significant differences for movements of the peripheral musculatures, was even more emphatic in respect of the disorganization of voluntarily sustained respiratory movement.

In addition to his experiments on control of the speech-mechanism, Blackburn carried out a subsidiary inquiry designed to compare the performance of stammerers and normal speakers in voluntarily maintaining regular movements of structures other than those of speech. His results revealed no significant difference between the two groups of subjects. Indeed, the most striking feature of Blackburn's tables is that all his subjects, stammering and normal alike, behaved in his tests of tapping-rate virtually like human metronomes,

showing in the great majority of the records zero standard deviations. It is hardly credible that human subjects could have maintained such a standard of regularity, and it may be questioned whether his method of measuring his kymographic records (which he does not specify) was not too crude to reveal the variations within and between the individual tracings. Since there are some apparent errors of computation in his presented data (Blackburn, 1931, table IV), it is not clear in what sense they should be taken. But it is certainly noteworthy that a reworking of his figures suggests a conclusion, although the differences are insufficient to substantiate it, that runs counter to the laterality theory that had been popularized by the Iowa laboratory where his work was done. So far as variation (i.e. erratic control) in tapping rate appeared at all in Blackburn's group, it was greater for the right hand than for the left in his normal subjects, and for the left hand than for the right in the stammering series.

II. The present experiments

Despite Blackburn's apparently negative results for tapping performance, there seemed to be some good reasons for investigating the matter more thoroughly than he had done. These were: (i) the intrinsic interest of the problem of 'sidedness', and the difficulty of arriving at a satisfactory test for the basic preference of the individual subject; (ii) the hypothesized relation between cerebral dominance, 'handedness', and the stammering disorder; and (iii) the suggestion, even in Blackburn's unsatisfactory data, of a difference between the stammerers and the normal speakers in respect of left- and right-hand control.

Consequently, in the present experiment the tapping test was run for both hands with two groups of subjects, consisting of fifteen ostensibly right-handed stammerers and fifteen right-handed normal controls. The subjects were children between the ages of 11 and 15; in each group there were two girls, and the groups were approximately matched for chronological age. The problem was to determine whether there are significant qualitative and quantitative differences in the unilateral motor control of stammerers and non-stammerers.

Kymographic records were obtained of three trials, each of approximately 15 sec. for each subject for each hand, giving an over-all recording time of 45 sec. for each hand. Subjects were required by the instructions to tap 'fairly quickly' on a tambour with a moderately thick rubber diaphragm on which was mounted a metal disk. During all trials the arm was unsupported, and right- and left-hand trials were taken alternately. The subjects were instructed to maintain, so far as possible, an even rate of tapping throughout the performance, but speed was not emphasized. The experimenter demonstrated what was required, and 15 sec. of unrecorded practice were allowed at the outset. The apparatus was concealed so that the subject could not regulate his performance by watching his own record.

Evaluation of the records

(1) Evaluation of the records required a count of the tapping movements recorded within a given time interval, fractional parts of a movement being assessed with reasonable accuracy by means of an opisometer. The section of the record under scrutiny was marked off by two vertical strands of fuse-wire stretched across a rectangular frame which could be moved along the time line like the cursor of a slide-rule. (2) In each 15 sec. record the first 3 sec. were ignored so that the subject might have had time to 'get into his stride' and establish a natural rhythm of performance. (3) The chief quantitative variable in which we are interested being *regularity* of performance, an appropriate sample of the subject's product was obtained by counting the movements for every third second of the record, time being recorded through-out in half-seconds.

The data for the individual subjects are set out in Table 1. In summary form, the relevant statistics from the two groups are as follows:

	Mean co-efficient of variation	
	Right hand	Left hand
Stammering group	9·39	6·17
Non-stammering group	4·87	5·16

Table 1. *Tapping records*

Subject	...	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Stammering group: (*a*) right hand																
M.		2·74	2·96	3·63	3·10	3·77	3·66	2·21	2·70	3·27	2·92	3·26	2·84	2·89	3·59	3·04
S.D.		0·37	0·42	0·15	0·28	0·33	0·50	0·15	0·20	0·17	0·27	0·41	0·28	0·17	0·39	0·27
V.		13·76	14·18	4·13	9·03	8·75	13·66	6·79	7·41	5·20	9·24	12·81	9·77	6·38	10·86	8·88
Stammering group: (*b*) left hand																
M.		2·54	2·85	3·39	2·92	3·31	3·10	2·24	2·14	3·14	2·92	3·14	2·72	2·66	3·14	2·93
S.D.		0·19	0·24	0·14	0·16	0·40	0·14	0·11	0·18	0·18	0·15	0·22	0·23	0·15	0·19	0·18
V.		7·31	8·42	4·13	5·48	12·08	4·52	4·91	8·41	5·73	5·13	7·16	8·34	5·94	6·05	6·20
Non-stammering group: (*a*) right hand																
M.		2·75	2·17	2·80	2·07	2·67	2·31	2·24	2·94	2·79	2·05	2·68	3·16	2·82	2·03	2·58
S.D.		0·24	0·14	0·10	0·04	0·14	0·12	0·16	0·12	0·11	0·09	0·11	0·15	0·11	0·08	0·14
V.		8·62	6·45	3·60	1·93	5·24	5·19	7·14	4·08	3·94	4·15	4·22	4·86	4·06	3·95	5·61
Non-stammering group: (*b*) left hand																
M.		2·45	2·43	2·73	2·11	1·90	2·97	2·01	3·16	2·25	2·24	2·21	2·78	2·53	1·97	2·46
S.D.		0·18	0·22	0·10	0·06	0·10	0·30	0·06	0·15	0·12	0·12	0·09	0·14	0·10	0·08	0·09
V.		7·22	9·05	3·62	2·84	5·26	10·10	2·99	4·74	5·33	5·35	4·08	5·02	4·13	4·06	3·66

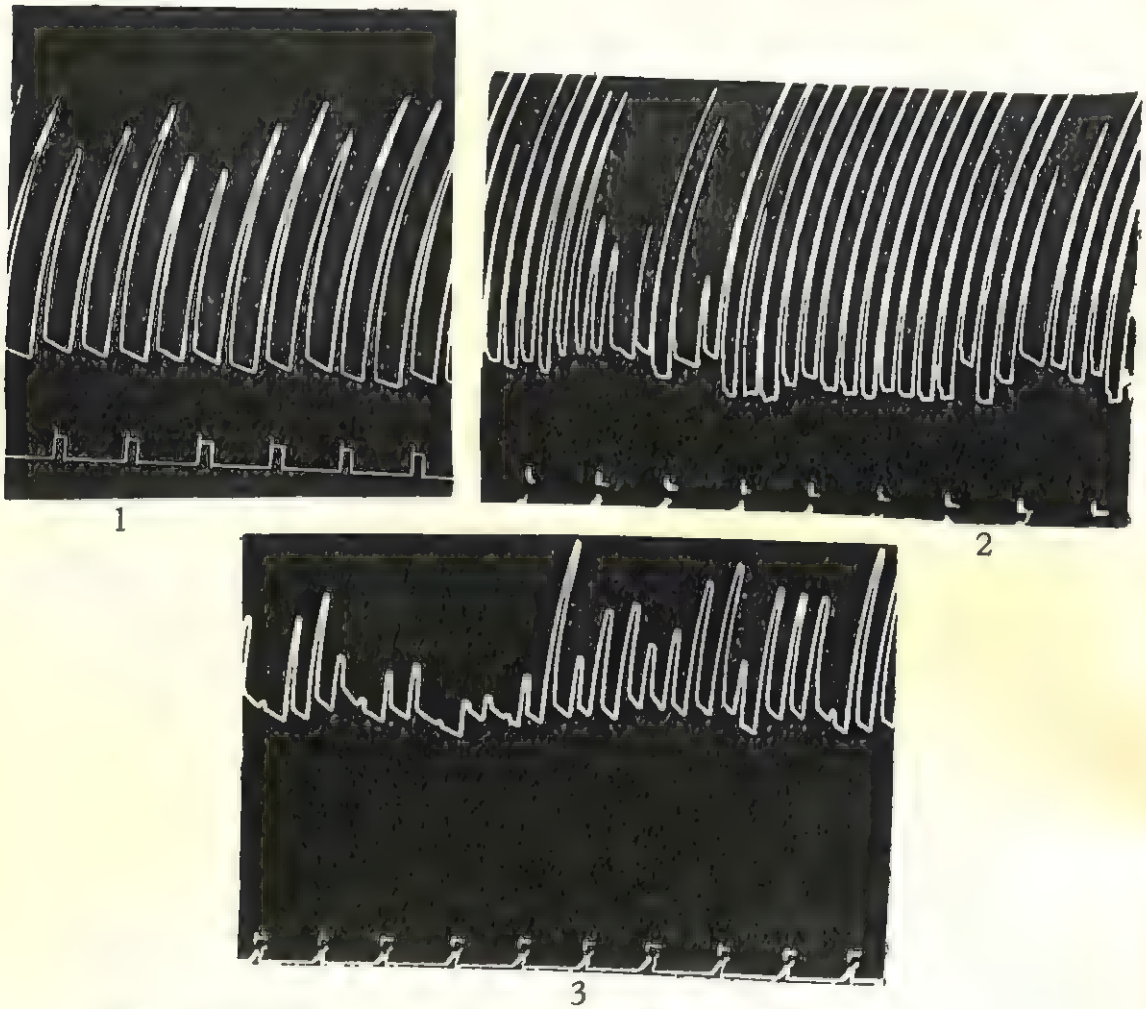
M. = average movements per second; V. = coefficient of variation (100 S.D./average).

III. DISCUSSION

Earlier experiments have shown that stammering subjects display a higher degree of variability in control of different parts of the speech musculature than do normal speakers. But some studies have tended to suggest (Seth, 1943) that this inferiority might even there be confined to the mechanism of respiration. The present result, however, makes it clear that, in our experimental group, the erratic control is not confined to the speech situation nor even to the speech mechanism, but extends to other forms of psychomotor performance as well. In particular, the following points should be noted. First, in the stammering group the coefficient of variation (i.e the inferiority in rhythmical performance) is very much greater for the apparently preferred right hand and arm than for the left; in the normal group, while there is no marked difference, the superiority lies in the preferred right hand. Using the Wilcoxon Matched Pairs Test, $p < 0{\cdot}01$ and $p > 0{\cdot}05$ respectively. Secondly, the coefficient of variation is greater for both hands in the stammering group, but it is very much greater when the right hands in particular are compared—9·39 as against 4·87. The Median Test gives: (*a*) right-hand, $p = 0{\cdot}0013$; (*b*) left-hand, $p = 0{\cdot}048$. Thirdly, examination of the individual results serves to emphasize both of these points. In only five cases, using the right hand, do the individual stammerers show a coefficient of variation lower than that of the most irregular of the control subjects, whereas for the left hand, in all but one of the stammering subjects, the records fall within the range of the normal group.

These results are clearly consistent with a neurophysiological theory of the stammering disorder on Orton–Travis lines. The disorganization of speech that we call stammering is certainly highly susceptible to changing features of the social situation in which speech is demanded; but then so also are the speech and other everyday motor performances of the patient suffering from cerebral lesion in the dominant hemisphere. It may be, therefore, that in the one condition as in the other the disorganization of nervous function must, in a proportion of the cases at least, be regarded as fundamental. Our results suggest, moreover, that stammering is not to be considered merely as a disorder of speech

but that, at least by early adolescence, the condition is demonstrable in some other aspects of psychomotor function, where it is ordinarily unsuspected. In so far as this implies a measure of consistency among the expressive acts of one stammering subject, it merits consideration in the light of the Allport–Vernon studies of expressive movement (Allport & Vernon, 1933). Their general conclusion is, indeed, so relevant to these results that it may be quoted directly. They say: 'Motor acts are not so specific as to be meaningless, and being organized they must reflect to a large degree the organization of the total brain-field. There are degrees of unity in movement, just as there are degrees of unity in



Figs. 1, 2 and 3. Tapping records. 1, normal speaker; 2 and 3, stammering speakers.

mental life and in personality. It is surely not unreasonable to assume that in so far as personality is organized, expressive movement is harmonious and self-consistent, and in so far as personality is unintegrated, expressive movement is self-contradictory.' Perhaps, however, the most important implication is this: that, while there is general agreement that the difficulty in the stammering disorder is central or cortical, our current attitudes, as they find expression in theory and in therapy alike, too often involve yet another instance of that 'neglect of the effectors' for which psychology in general has

been rightly criticized, and more energetic pursuit of inquiry into a variety of aspects of psychomotor function in a wider range of groups, especially of so-called 'functional' or organic disturbance, might be rewarding.

Some typical records are reproduced which serve to illustrate the foregoing points. These may be compared with those presented for respiration, etc., in the earlier studies (Blackburn, 1931; Seth, 1934) to which reference has been made. They provide concrete and striking qualitative evidence for the points that are summarized and generalized in the tabulated data. Fig. 1 is taken from the middle of the third record of tapping with the index finger of the right hand, produced by the most regular of the non-stammering subjects (no. 14, S.D. 0·08). The quantitative and qualitative uniformity of the record is particularly striking, and the regularity in the height of the separate peaks of the tracing, which answers to the force with which each of the separate taps was delivered, is noteworthy. A corresponding record by a stammering subject (no. 1, S.D. 0·37; third trial with the right hand) is presented in Fig. 2, the inferiority of the control being apparent in the general arhythmical 'jerkiness' of the tapping and the wide range of variation in the force of the percussion. Qualitative irregularity is even more obvious in Fig. 3, in which is presented part of the performance of another stammering subject (no. 14, third trial with the left hand) whose quantitative record is not by any means unique (S.D. 0·19). But the standard deviation, depending upon count alone, perforce must disregard, and in some cases obscure, the haphazard manner in which the tracing (i.e. the forcefulness of the tapping) fluctuates from moment to moment, with no appearance of periodicity.

REFERENCES

ALLPORT, G. W. & VERNON, P. E. (1933). *Studies in Expressive Movement.* New York: Macmillan.

BLACKBURN, B. (1931). Voluntary movements of the organs of speech in stutterers and non-stutterers. *Psychol. Monogr.* **41**, 4.

BURT, C. (1937). *The Backward Child.* London: University of London Press.

SETH, G. (1934). An experimental study of the control of the mechanism of speech in stuttering subjects. *Brit. J. Psychol.* **24**, 375–88.

TRAVIS, L. E. (1931). *Speech Pathology.* New York: Appleton and Co.

(*Manuscript received* 8 *April* 1957)

COGNITIVE AND SOCIAL CHANGES IN THE FEEBLEMINDED —THREE FURTHER STUDIES

By A. D. B. CLARKE, A. M. CLARKE and S. REIMAN

Psychology Department, The Manor, Epsom, Surrey

Earlier papers have described researches showing that I.Q. increments in adolescent and young adult persons certified as feebleminded occur not infrequently, cannot be explained as artefacts, and appear to be linked with early very adverse environmental experiences. The present three studies indicate that I.Q. increments can be large over long periods of time, are correlated with changes in social adjustment, but cannot be shown to be induced by special environmental stimulation; in the present age range and type of subject, they tend to occur in any reasonable environment. Results imply that a necessarily poor outcome cannot be predicted for children with I.Q.'s in the 50's, 60's or 70's if they also come from very adverse conditions. Test reliability, statistical regression, and a review of some relevant literature are discussed; theoretical problems concerning intellectual development are presented; and attention is directed to individual differences in susceptibility to environmental damage and to the whole problem of human resilience.

I. Introduction

In earlier papers (Clarke & Clarke, 1953, 1954*b*, 1955) the general problem of I.Q. inconstancy was considered and, specifically, the variability of the I.Q. among adolescents and young adults certified as feebleminded was discussed. A two-year study indicated that I.Q. increments were of frequent occurrence, particularly in cases characterized by extreme deprivation in early life. In a population drawn, on the whole, from very adverse socio-economic conditions, those from the worst had made an average increment of 9·7 points, while a residual group from somewhat less unfavourable conditions made an average increase of 4·1 points. This latter was similar in amount, though not necessarily of the same nature, as the gains arising from the maximal effects of test practice and errors of underestimation in a control group retested after a short time interval. Cognitive changes in the certified feebleminded differed from those occurring in normals in that: (1) they tended to be mainly in the direction of improvement; (2) in relation to time interval between test and retest, they tended to be rather large; and (3) they took place at ages when intellectual growth is commonly assumed to have ceased. The general picture which emerged from this and other studies was that many high-grade defectives showed temporary arrest of intellectual growth, while later reverting towards (and sometimes achieving) normality.

When the earlier work (which throughout the present paper will be termed the *original research*) was completed, it was decided to plan longer-term investigations which might throw further light on the determinants and limits of these increments, and which might have further implications for our understanding of mental deficiency. Objective psychometric assessments have been carried out at this hospital since 1949, and it was considered that by 1955–6 sufficient data would be available to make such longer-term research possible. In all, three new studies are reported here.

II. First Study

Cognitive changes

The first experiment was designed to follow-up those members of the *original research* who were still available. The original subjects had, in 1952, been allotted either to a group from very bad homes (based on the presence of two or more of twelve criteria* in the case history) or to a residual group from less adverse circumstances; this rating had been carried out independently by an external investigator who knew neither the patients nor their test scores. It seemed clear that the 28 available at the end of 1955 (from an original 59) were, in the main, the less satisfactory members of the sample, since a large proportion

Table 1. *Test-retest data, Wechsler, Form I,† over 6 years*

	1949, Full Scale I.Q.	1952, Full Scale I.Q.	1955, Full Scale I.Q.	Diff. '49/'52	Diff. '52/'55	Diff. '49/'55
			(*a*) Group from very bad homes			
Mean	59·6 (66·4)	70·7 (76·1)	75·8	11·1 (9·7)	5·1	16·2
S.D.	9·7 (14·1)	11·4 (13·4)	10·3	4·2 (6·3)	5·0	6·1
N	9 (25)	9 (25)	9	9 (25)	9	9
t (1-tail test)	—	—	—	7·929	3·059	7·969
Significant at	—	—	—	0·1%	1%	0·1%
			(*b*) Group from less adverse homes			
Mean	62·3 (66·1)	66·8 (70·1)	72·5	4·5 (4·1)	5·7	10·2
S.D.	13·4 (14·1)	13·0 (13·1)	13·3	4·4 (4·9)	5·1	6·6
N	19 (34)	19 (34)	19	19 (34)	19	19
t (1-tail test)	—	—	—	4·460	4·872	6·737
Significant at	—	—	—	0·1%	0·1%	0·1%

I.Q. range: 1949, 47–72; 1952, 58–86; 1955, 64–91.
Age at final retest: mean 26·4, S.D. 5·5 years.
Period between first and final test: mean 70·9 months, S.D. 2·5 months.
Range of I.Q. changes: +3 to +22 points. (10 points and above: 8/9 = 89%; 15 points and above: 7/9 = 78%.)

I.Q. range: 1949, 35–87; 1952, 40–89; 1955, 48–100.
Age at final retest: mean 27·9, S.D. 9·5 years.
Period between first and final test: mean 74·2 months, S.D. 5·3 months.
Range of I.Q. changes: −2 to +21 I.Q. points. (10 points and above: 10/19 = 53%; 15 points and above: 5/19 = 26%.)
t (between group from very bad homes and group from less adverse homes) = 2·403 (significant above 2·5% level, 1-tail test).

(54%) of the remainder had by now been discharged either as being 'no longer mentally deficient within the meaning of the Act' or as being sufficiently competent socially to warrant this action, or alternatively, were on distant licence and likely to achieve discharge in the near future. Table 1 (*a*) and (*b*) indicate that the present subjects were, both in 1949 and in 1952, on the average less intelligent than the remainder of the sample

* The main point in each of these criteria was as follows: N.S.P.C.C. intervention, parental attitude antagonistic, no fixed abode, 'Fit Person' Order, home conditions bad, considerable neglect, irregular school attendance due to neglect, home dirty and neglected, gross poverty, crime in parents, rickets, child found begging.

† Fifty patients drawn from the present three studies were, in addition, given the Progressive Matrices test to confirm their present levels. These correlated to the extent of 0·81, with a mean difference of 2·67 points between them in favour of the Wechsler.

COGNITIVE AND SOCIAL CHANGES IN THE FEEBLEMINDED —THREE FURTHER STUDIES

By A. D. B. CLARKE, A. M. CLARKE and S. REIMAN

Psychology Department, The Manor, Epsom, Surrey

Earlier papers have described researches showing that I.Q. increments in adolescent and young adult persons certified as feebleminded occur not infrequently, cannot be explained as artefacts, and appear to be linked with early very adverse environmental experiences. The present three studies indicate that I.Q. increments can be large over long periods of time, are correlated with changes in social adjustment, but cannot be shown to be induced by special environmental stimulation; in the present age range and type of subject, they tend to occur in any reasonable environment. Results imply that a necessarily poor outcome cannot be predicted for children with I.Q.'s in the 50's, 60's or 70's if they also come from very adverse conditions. Test reliability, statistical regression, and a review of some relevant literature are discussed; theoretical problems concerning intellectual development are presented; and attention is directed to individual differences in susceptibility to environmental damage and to the whole problem of human resilience.

I. Introduction

In earlier papers (Clarke & Clarke, 1953, 1954*b*, 1955) the general problem of I.Q. inconstancy was considered and, specifically, the variability of the I.Q. among adolescents and young adults certified as feebleminded was discussed. A two-year study indicated that I.Q. increments were of frequent occurrence, particularly in cases characterized by extreme deprivation in early life. In a population drawn, on the whole, from very adverse socio-economic conditions, those from the worst had made an average increment of 9·7 points, while a residual group from somewhat less unfavourable conditions made an average increase of 4·1 points. This latter was similar in amount, though not necessarily of the same nature, as the gains arising from the maximal effects of test practice and errors of underestimation in a control group retested after a short time interval. Cognitive changes in the certified feebleminded differed from those occurring in normals in that: (1) they tended to be mainly in the direction of improvement; (2) in relation to time interval between test and retest, they tended to be rather large; and (3) they took place at ages when intellectual growth is commonly assumed to have ceased. The general picture which emerged from this and other studies was that many high-grade defectives showed temporary arrest of intellectual growth, while later reverting towards (and sometimes achieving) normality.

When the earlier work (which throughout the present paper will be termed the *original research*) was completed, it was decided to plan longer-term investigations which might throw further light on the determinants and limits of these increments, and which might have further implications for our understanding of mental deficiency. Objective psychometric assessments have been carried out at this hospital since 1949, and it was considered that by 1955 6 sufficient data would be available to make such longer-term research possible. In all, three new studies are reported here.

II. First Study

Cognitive changes

The first experiment was designed to follow-up those members of the *original research* who were still available. The original subjects had, in 1952, been allotted either to a group from very bad homes (based on the presence of two or more of twelve criteria* in the case history) or to a residual group from less adverse circumstances; this rating had been carried out independently by an external investigator who knew neither the patients nor their test scores. It seemed clear that the 28 available at the end of 1955 (from an original 59) were, in the main, the less satisfactory members of the sample, since a large proportion

Table 1. *Test-retest data, Wechsler, Form I,† over 6 years*

	1949, Full Scale I.Q.	1952, Full Scale I.Q.	1955, Full Scale I.Q.	Diff. '49/'52	Diff. '52/'55	Diff. '49/'55
			(*a*) Group from very bad homes			
Mean	59·6 (66·4)	70·7 (76·1)	75·8	11·1 (9·7)	5·1	16·2
S.D.	9·7 (14·1)	11·4 (13·4)	10·3	4·2 (6·3)	5·0	6·1
N	9 (25)	9 (25)	9	9 (25)	9	9
t (1-tail test)	—	—	—	7·929	3·059	7·969
Significant at	—	—	—	0·1%	1%	0·1%

I.Q. range: 1949, 47–72; 1952, 58–86; 1955, 64–91.
Age at final retest: mean 26·4, S.D. 5·5 years.
Period between first and final test: mean 70·9 months, S.D. 2·5 months.
Range of I.Q. changes: +3 to +22 points. (10 points and above: 8/9 = 89%; 15 points and above: 7/9 = 78%.)

			(*b*) Group from less adverse homes			
Mean	62·3 (66·1)	66·8 (70·1)	72·5	4·5 (4·1)	5·7	10·2
S.D.	13·4 (14·1)	13·0 (13·1)	13·3	4·4 (4·9)	5·1	6·6
N	19 (34)	19 (34)	19	19 (34)	19	19
t (1-tail test)	—	—	—	4·460	4·872	6·737
Significant at	—	—	—	0·1%	0·1%	0·1%

I.Q. range: 1949, 35–87; 1952, 40–89; 1955, 48–100.
Age at final retest: mean 27·9, S.D. 9·5 years.
Period between first and final test: mean 74·2 months, S.D. 5·3 months.
Range of I.Q. changes: −2 to +21 I.Q. points. (10 points and above: 10/19 = 53%; 15 points and above: 5/19 = 26%.)
t (between group from very bad homes and group from less adverse homes) = 2·403 (significant above 2·5% level, 1-tail test).

(54%) of the remainder had by now been discharged either as being 'no longer mentally deficient within the meaning of the Act' or as being sufficiently competent socially to warrant this action, or alternatively, were on distant licence and likely to achieve discharge in the near future. Table 1 (*a*) and (*b*) indicate that the present subjects were, both in 1949 and in 1952, on the average less intelligent than the remainder of the sample

* The main point in each of these criteria was as follows: N.S.P.C.C. intervention, parental attitude antagonistic, no fixed abode, 'Fit Person' Order, home conditions bad, considerable neglect, irregular school attendance due to neglect, home dirty and neglected, gross poverty, crime in parents, rickets, child found begging.

† Fifty patients drawn from the present three studies were, in addition, given the Progressive Matrices test to confirm their present levels. These correlated to the extent of 0·81, with a mean difference of 2·67 points between them in favour of the Wechsler.

now no longer available, and this fact adds support to the belief that these were indeed the less satisfactory members of the original sample. Thus, the results of a follow-up of those remaining would be expected to give a minimal estimate of the I.Q. change over a period of 6 years for patients of this age, type and social background. Table 1(*a*) and (*b*) show the test and retest data on Wechsler, Form I, for the members of the two groups. The figures in brackets indicate the relevant data for the original groups of which the present subjects formed a part.

It will be noted that the I.Q.'s of those from very bad homes present a substantially different cognitive picture in 1955 from that in 1949, having, on the average, increased by about one standard deviation. The I.Q.'s of those from less adverse home circumstances had changed by about two-thirds of a standard deviation, and it is clear that in their case the change had been slower, smaller in amount, and had been roughly equal in the two equivalent time periods, 1949–52 and 1952–5. Members of the former group, however, had shown the maximum change in the earlier part of this period, thereafter improving less rapidly. This does not conflict with the earlier hypothesis that I.Q. increments seemed to be the effect of a fading of 'intellectual scars' induced by bad home conditions, and occurred more as a result of *removal from* such conditions rather than of entry into relatively better ones. Those with the greater damage due to deprivation made the greater recovery (often many years later) and they made it more rapidly.

Social changes

A significant social factor is reflected by the difference between the two groups in terms of present availability; at the end of 1955 only nine of the original twenty-five from very bad homes were still available, whereas a larger proportion, nineteen out of thirty-four, from less adverse conditions were still under the care of the hospital. Since many of the remainder had been discharged, this adds support to the generalization that, in this type of population, the worse the social background, the better the prognosis from all points of view. Moreover, of the nine remaining from the original group from very bad homes, eight were already working in the community although still under certificate, and one was receiving industrial training prior to community placement. Only four of the nineteen remaining from the original group from less adverse circumstances were in a similar position at the end of 1955.

It is unlikely that I.Q. increments play a major part in *causing* improved social adaptation, since the correlation between I.Q. and social competence at this level is not very high. These increases must therefore be regarded as one facet of a process of more or less total personality development which is also likely to be associated with the same general causation.

III. Second Study

Cognitive changes

The second study followed up all patients of feebleminded grade first admitted to this hospital in 1951 who were still available at the beginning of 1956. The purpose of this investigation was to find out whether, among other things, I.Q. changes in this population and age range were a direct function of time interval between test and retest, as is sometimes implied for normals (*cf.* Eysenck, 1953). Retesting these subjects in 1955–6 gave

an interval between tests of 4½ years, roughly double that of the *original research*; it was thus possible to ascertain whether the average increments would also be doubled. Of an original sixty patients admitted in 1951, thirty-two were still available, and, as before, it seemed likely that they were the less promising members of the original sample.

The following methodology was employed. An independent investigator from another hospital examined in detail the thirty-two case-histories, using the twelve criteria evolved in the *original research* (see footnote, p. 145) and divided them into a group from very adverse homes and a residual group from less adverse conditions. As before, this latter contained many from bad homes, though not exceptionally bad, and very few from good homes. The external investigator knew neither the patients nor their test scores. It was then predicted that members of the former group would tend to make the larger increments. Subsequently, the retesting was carried out by one of us (S.R.) who was ignorant of the patients, their social histories, the group to which they had been allotted, the prediction which had been made, and their earlier test scores. The results are shown in Table 2.

Table 2. *Patients admitted in* 1951, *retested* 1955–6

	Group from very bad homes			Group from less adverse homes		
	Wechsler Full Scale I.Q., 1951	Wechsler Full Scale I.Q., '55/'56	Diff.	Wechsler Full Scale I.Q., 1951	Wechsler Full Scale I.Q., '55/'56	Diff.
Mean	63·1	74·6	11·5	69·0	76·2	7·2
S.D.	15·0	15·2	6·3	13·9	14·6	6·6
N	13	13	13	19	19	19
t (1-tail test)	—	—	6·894	—	—	4·926
Significant at	—	—	0·1%	—	—	0·1%
I.Q. ranges:	1951, 37–90; 1955–6, 54–99			1951, 42–94; 1955–6, 55–101		
Age at retest:	Mean 23·0, S.D. 5·9 years			Mean 23·6, S.D. 6·1 years		
Period between test and retest:	Mean 53·6 months, S.D. 3·8 months			Mean 51·6 months, S.D. 5·0 months		
Range of I.Q. changes:	0 to +18 points. (10 points and above: 9/13 = 69%; 15 points and above: 7/13 = 54%.)			−7 to +20 points. (10 points and above: 8/19 = 42%; 15 points and above: 1/19 = 5%.)		

t (between group from very bad homes and group from less adverse homes) = 1·925, significant at the 5% level (1-tail test). This difference is much less significant than in the *original research*, and probably arises from a combination of four factors: (1) Total *N* is much smaller. (2) One patient in the group from less adverse homes showed a very atypical gain of 20 points, which alone was responsible for depressing statistical significance from the 2½% to the 5% level; at the time of first testing it had been noted that she was unusually disturbed emotionally, and came from a background, which though not quite conforming to the necessary criteria, was far from adequate. (3) The only patient in the group from very bad homes who did not change at all was later found to have been suffering from a right temporal malignant tumour during the material period. (4) Selective wastage had almost certainly occurred since 1951, since rehabilitation and discharge rates have accelerated greatly since then. Consequently, there has probably been a tendency for the better members of the group from very bad homes to have left the institution.

It will be seen that once again the subjects from very bad homes tended on the average to make the larger increments, although in this sample the discrepancy between the *average* increments by the two groups is less marked than in the *original research*. With roughly double the original time interval, it is clear that, on the one hand, the present group from very bad homes did not increase greatly their average increment compared with that achieved by the original group (mean increment 11·5 compared with 9·7 points). Clearly, therefore, the amount of gain in this group is not a direct function of the length

of time between test and retest. On the other hand, it will be seen that members of the residual group from less adverse conditions almost doubled their mean increment over the doubled time interval (mean increment of 7·2 compared with 4·1 in the *original research*). Thus their gains are much more a direct function of the test-retest time interval. This is consistent with findings in the present *First Study* in which the maximal gain for the group from very bad homes occurred in the first 2–3 years, with a relatively slow increase thereafter, while those from less adverse conditions made a steady increase over the 6-year period. Thus, the difference between the gains made by both the groups is maximal at the beginning of the period in which increments have been studied, and subsequently because the growth rate of the bad home group tends to decelerate sharply (see Table 1), while the residual group maintains a steady progress, the difference between the average gains decreases.

Social changes

The present sample on the average had been under care of this hospital a shorter time than any of the others and was younger. Nevertheless, social changes had already begun to be apparent, and differences, as yet small, between the two groups were noticeable. Thus seven out of the thirteen from very bad homes were either undergoing intensive training prior to community placement (see next Study) or were already in outside employment, while only six out of the nineteen from the less adverse homes were similarly placed.

IV. Third Study

This study was devoted to a special problem; the *original research* had suggested that I.Q. changes resulted as an effect of removal from very adverse conditions rather than from entry into relatively better ones; that is, absence of such unfavourable circumstances enabled intellectual scars to fade. An important problem then envisaged was to determine whether special environmental conditions would aid this apparently natural process of recovery. A group of patients was readily available for the testing of this hypothesis.

For some years the Psychology Department has been responsible for running two small rehabilitation units for selected high-grade patients. These are taught industrial skills in small workshops. develop good work habits, receive educational aid and experience of the community, eventually being placed in local industry at standard wages and in normal jobs. During their initial period outside they return to the hospital nightly but after a time may graduate to lodgings and obtain their discharge. Observation, and their success rate, suggested that many undergo radical changes under these conditions, and the question thus became whether these involved intellectual aspects of personality, and, if so, whether these gains were environmentally induced. The scheme briefly mentioned has been described in detail elsewhere (1954*a*).

At any one time about twenty to twenty-five patients so trained are at work in the community, and these formed the subjects of the present experiment. All had at some time during the last few years been given Wechsler. Form I, and this enabled retest results to be compared with their earlier scores. During the period between test and retest all had had at least 6 months intensive training, and at least 6 months in the community (and often more), both of which involved considerable change of circumstances for these persons. Social histories before admission were again assessed by the external investigator

who knew neither the patients nor their test scores, and retesting was carried out by one of us (S.R.) who knew neither the patients, nor their early histories, nor their earlier test scores. Table 3 shows the relevant data for both groups.

The first point of note is the rather higher proportion than usual of the whole group (a little over half, compared with about 42% in a more random sample) who came from very bad homes; this is likely to be the result of a selection factor, namely that such persons tend to do better socially than others and are therefore more likely to be chosen for special training and rehabilitation. Once again, the larger increments tended to occur in these patients, although the difference did not reach an acceptable level of significance; this may have resulted from a selection factor favouring the more promising patients

Table 3. *Wechsler full-scale test-retest results on persons specially trained in rehabilitation units and placed in the community in normal employment*

	Group from very bad homes			Group from less adverse homes		
	Test I.Q.	Retest I.Q.	Diff.	Test I.Q.	Retest I.Q.	Diff.
Mean	64·8	78·9	14·1	75·1	85·5	10·4
S.D.	13·7	13·1	6·0	13·4	12·4	6·3
N	11	11	11	10	10	10
t (1-tail test)	—	—	7·772	—	—	5·547
Significant at	—	—	0·1%	—	—	0·1%
I.Q. ranges:	First test, 49–94; Retest, 64–105			First test, 63–108; Retest, 68–112		
Age at retest:	Mean 24·2, S.D. 3·9 years			Mean 27·6, S.D. 6·9 years		
Period between test and retest:	Mean 54·6 months, S.D. 17·6 months			Mean 47·0 months, S.D. 17·6 months		
Range of I.Q. changes:	+5 to +20 I.Q. points. (10 points and above: 9/11 = 82%; 15 points and above: 7/11 = 64%.)			+2 to +19 I.Q. points (10 points and above: 5/10 = 50%; 15 points and above: 3/10 = 30%.)		
Time in training:	Mean 10·6 months, S.D. 6·2 months			Mean 12·4 months, S.D. 5·2 months		
Time in employment:	Mean 13·2 months, S.D. 10·4 months			Mean 16·4 months, S.D. 11·7 months		

t (between group from very bad homes and group from less adverse homes) = 1·449, not significant (1-tail test).

from the less adverse conditions. While in both groups the average increase is rather large, it is not so very different from what would be expected under ordinary hospital conditions. We are able to compare these data with the results of the *Second Study*, Table 2. The average test-retest time interval is similar, and in the *Second Study* the average gain of those from very bad homes was 11·5 points, S.D. 6·3, while in the present study the similar group increased on the average by 14·1 points, S.D. 6·0. Similarly, the group from less adverse homes in the *Second Study* made an increment of 7·2 points, S.D. 6·6, compared with an average of 10·4, S.D. 6·3, from the similar group in the present investigation. The average difference between the gains in the *Second* and *Third Studies* is thus a rough measure of the effect of special environmental conditions; for the groups from bad homes the difference is 14·1 minus 11·5 = 2·6 points, and for the remaining group 10·4 minus 7·2 = 3·2 points. These are then the maximum effects which can be ascribed to environmental difference, and they are clearly small.

As has been stated, subsequent I.Q. increment seems to occur as a long-term effect of removal from a traumatic environment, and there can be no doubt that the earlier intellectual deficit resulted, partly at least, from *negative* environmental effects in childhood. The present study has indicated that fairly clear environmental differences in this age-group do not exert much of a *positive* effect upon the I.Q. There are two possible explanatory factors. First, the difference between the ordinary institutional environment

and the one provided by intensive training in special units followed by community experience may not be as large as thought. On the other hand, it must be noted that the difference is certainly large enough to produce considerable social changes, resulting in greater success and more rapid discharge, so that there is little doubt that the environment is a specially stimulating one.

Secondly, it is possible that the findings reflect both an initial *negative* effect in childhood and a *masked positive* effect in adolescence and early adult life. It may be that any environment, other than an adverse one, will stimulate the apparently natural recovery process, and that the effects of environmental differences are small simply because all—even the institutional environment—have a *positive* effect. That environmental effects cannot themselves be measured needs stressing; only the effects of environmental differences are measurable. Precisely this same point has been made about the classic studies of identical twins brought up apart; in only a few cases was each twin in a radically different environment and thus the variable 'environment' was seldom accessible for full measurement.

Another source of evidence bearing out the present conclusions emerges from a comparison of the results of the *First Study* with this one. The time intervals are not quite similar, but it will be noted that the members of both groups from very bad homes made very similar mean increments. Eight of the nine members of this group in the *First Study* were employed in the community in jobs traditionally provided for the mentally deficient and for which they had no special training; yet their I.Q. gains were closely similar to those in the present study. So, too, with the members of both groups from less adverse homes; those who had had no special training and the majority of whom had had no experience of outside employment (*First Study*) made similar increments to those who had had both. There was no doubt, however, that special training and normal experiences had induced profound social changes in the patients of the present study. They were all holding good jobs with wages ranging from £3. 0*s*. 0*d*. to £8. 2*s*. 0*d*. per week with an average of £5. 17*s*. 0*d*. All were expected to gain their discharge, and to become reasonable and self-supporting citizens.

Summing up, it seems clear that intellectual retardation or arrest can and does occur in conditions really antagonistic to mental health in childhood. And it is precisely these conditions from which the vast majority of the certified feebleminded are drawn. Entry into any better environment seems equally able to stimulate increments in I.Q. and to allow the damage to fade. This, however, is a long-term process which will be discussed in detail in the next section.

V. Discussion

(*a*) *Variability of the* I.Q.

At this stage, brief discussion of the causes of I.Q. variability is appropriate. Spurious changes can be subsumed under the first three of the following headings. (1) *Errors of measurement* which in most cases represent *personal fluctuations* (e.g. alterations in mood, misunderstanding of the instructions, chance decisions, fatigue, boredom, anxiety and changes in motivation on the part of the subject, and errors in scoring, instructions, or in the establishment of rapport on the part of the tester). These combine to reduce test-retest correlation from a theoretical +1·0 to an average of 0·9 on immediate retest

(Thorndike, 1933). Similarly, standardization differences (e.g. varying S.D.'s on the 1937 Stanford–Binet at different ages) will, if uncorrected, produce spurious changes. (2) *Test practice and coaching* have well-known effects (see Vernon, 1954). (3) *Incorrect testing* by persons unqualified in psychometrics may clearly produce errors resulting in excessive I.Q. variability upon retest. (4) *Variation in the rate of intellectual growth.* In general, as the interval between tests lengthens, so the retest correlation decreases, often steadily (Nemzek, 1933; Thorndike, 1940; Eysenck, 1953). There is no reason to suppose that errors of measurement (whether arising from personal fluctuations in respect of subject or tester, or standardization errors) are likely to be steadily cumulative, and thus a decreasing correlation coefficient must be primarily a reflexion of some real change of status of members of the group.

In this connexion, it is necessary to discuss 'test reliability' since confusion tends to arise over this concept. Test reliability relates to two factors: first, whether the test measures the same quality in the same way on different occasions; and secondly, whether the quality measured itself alters. The first point has been discussed under 'errors of measurement', and the second might well be termed 'unreliability of persons' since, as we have seen, the fact of decreasing correlation with increase of time implies just that. It is this in particular which produces difficulties; it is by no means uncommon to hear or read the following tautological argument: 'over several years test results change because of imperfect correlation or test unreliability'. It is further implied that such changes are statistical, reside in the test rather than in the person tested, and therefore have no psychological significance. Obviously, however, the test unreliability or lowered correlation over several years is more likely to be a reflexion of change of status of individuals rather than vice versa.

Statistical regression towards the mean is another concept frequently used to account for I.Q. changes; where this concerns errors of measurement on the same or different tests, or where different tests have a low initial correlation when given within a day or two of each other, then this is entirely correct. In such cases, regression of extreme scores is inevitable and of no psychological significance. In a cogent discussion of regression, McNemar (1940), agreeing that it is a descriptive rather than an explanatory concept, points out that an inferior or superior group will not move towards the general population mean on retest unless it has been specifically selected on the basis of initial test results (e.g. Terman's group of gifted children, selected on the basis of I.Q. 140 and above). Within a group not so selected, regression effects will occur around the group mean, be it superior or inferior. We have already seen that, granted proper standardization and test reliability (i.e. small errors of measurement) a steadily decreasing correlation coefficient tends to occur with increase of time, such that after 5 years it may be about 0·7. This implies some changes of as much as 15–20 points, and it would be quite incorrect to say that these are 'due to regression' with consequently no psychological meaning. Yet such statements are not uncommonly made, and in effect there appears to be a tendency to regard statistical expressions as somehow primary to behavioural data, with curious results. Regression cannot explain such large changes which reflect primarily an alteration in response to the test situation; naturally the implications of such variations are not necessarily obvious and must be interpreted with caution.

At the risk of labouring this point, let us see why in the *original research* and the present studies, the concept of regression cannot explain the major part of the increments.

First, the groups were not chosen for study or follow up on the basis of initial test score. Secondly, the *original control group* retested after a short time interval made an average increment of only 4 points, which must be a maximal estimate of errors of measurement including regression effects for this type of sample. Stating the same point differently, the Wechsler test has a very high immediate retest reliability, hence regression effects must be small. Thirdly, such wastage of subjects as occurred in the *First* and *Second Studies* was for reasons other than initial test results. Fourthly, the two groups which formed the main sample in the *original research* were initially equal in I.Q. Yet by the time that investigation was complete, and even more so now, different cognitive and social trends had occurred, and regression cannot account for such differences in initially equated groups. Fifthly, statistical regression is psychologically meaningless. Throughout all our studies, however, there has been a consistent tendency for the amount of 'regression' to be related to psychologically meaningful external criteria, the record of early experiences. It is hardly likely that the purely statistical would be so linked with the purely behavioural. Thus on all counts, statistical regression cannot explain the major portion of the I.Q. increments.

(*b*) *The present findings*

The increments noted both in the *original research* and in the present studies are likely to have been largely the effect of 'unreliability of persons', that is, of variations in the rate of late intellectual maturation. All four investigations underline the better prognosis for the feebleminded person from bad social conditions; they tend to have larger I.Q. increments, to enjoy a greater success within the hospital, to be selected for intensive training by virtue of their other qualities, to respond to training a little more rapidly and to be tried in the community earlier and more often than those from rather less adverse social conditions. They also tend to achieve discharge from care more rapidly and more often, and although sometimes they may have a lower I.Q. in adolescence than the other patients, they tend to catch up and overtake them within a few years. Indeed they present a substantially different picture in many spheres after, say, 6 years, and often these changes occur at relatively late ages. In general, these results are strongly supported by the social findings of Charles (1953) in which a very long-term follow-up, conducted with exemplary care, of children originally judged mentally deficient, showed that they had not developed according to a rigid stereotype. Many at the ages of 30 and above were in the normal or dull normal range and were holding their own in the community. The outcome for the majority was far better than would have been predicted at the time of their attendance in the 'opportunity rooms' of Lincoln, Nebraska, schools. Similarly, any person reading the case-histories of many of the present subjects cannot fail to be struck by the discrepancy between the gloomy predictions of psychologists and psychiatrists made within the last 10 years, and the actual progress made. Delayed personality maturation may well be a factor explanatory of Penrose's (1949) observation that the incidence of high-grade deficiency decreases sharply from the age of 15 onwards.

What, then, are the precise mechanisms of these changes? Earlier papers (1954*b*,1955) argued that at all stages the obtained I.Q.'s were relatively accurate measures of the then intellectual status of the subjects, rejecting the hypothesis of pseudo-feeblemindedness, which, as will be seen, really begs the question. The work of Hebb (1949) is clearly relevant; adverse conditions in early life would be likely to interrupt or seriously damage the for-

mation of 'phase sequences', and the more adverse these conditions the greater the deficit to be made good. It appears that improvement may not commence until adolescence and may continue well into the twenties. In many cases there thus appears to be a period of latency, during which growth rates may show little acceleration, lasting from childhood and infancy when the damage occurred, until adolescence. We possess far less data on this period, however, and this view is inferred from several sources: (*a*) the few children studied with similar backgrounds to those of our older subjects appear to change little; (*b*) the case-histories of our subjects do not imply, in most cases, a level in childhood very much inferior to their status in adolescence; and (*c*) most of the early investigators believed that intellectual growth in the feebleminded ceased several years earlier than in the normal. Our own results suggest precisely the opposite for many such persons, but it is a reasonable assumption that this original view about intellectual development in the feebleminded sprang from the fact that children were mainly being studied at that time, and it may be that over late childhood there is a developmental plateau which, when it begins, suggests a cessation of growth. While this latency period is an hypothesis, it seems to be a reasonable one, and at all events it seems quite certain that in childhood these persons do not change as rapidly as later. Removal of most feebleminded children from traumatic conditions usually took place years before entry into a mental deficiency hospital, the interim period being spent in residential E.S.N. schools, children's homes or orphanages. Then, perhaps at age 16, certification took place as the last link in a long chain of official intervention.

All one can say in explanation of this hypothetical latency period is that it may be needed to make good the foundations upon which subsequent development depends, and to establish or re-establish some integration of personality—that perhaps 5 or 10 years are needed to give the developmental basis normally gained during the first few years of life. Certainly future investigations must be concerned with studying such persons at a much earlier age, and as soon as possible after removal from adverse homes.

A further and necessary hypothesis requiring some discussion is that there must be limits to the change which occurs; these limits may be set by the basic resources of the central nervous system, genetically determined, and also, perhaps, by the amount and quality of the damage produced by negative environmental influence. The largest increment ever noted in this hospital on the same test was 29 points but this was over about 7 years; this then appears to be the maximum over a similar time period in adolescence and early adult life among such subjects. Over the whole of a person's life span, however, the limits to change may be slightly greater, say, a maximum of 40 points (cf. Vernon, 1955); obviously, however, the majority of the feebleminded may never make changes of this order. Not only do the amounts of change differ in different subjects but also the rates of change. We have seen in the *First Study* that those from really adverse conditions tended to make a large initial gain over about 3 years, which on the average was halved during the next 3 years; those from less adverse homes (but nevertheless bad ones) tended to make a steady gain over the 6-year period. The first interpretation mentioned was that the rate of recovery depended upon the severity of initial damage. The concept of limits may, however, also provide an explanation. It may be that the members of the group from very bad homes were reaching their limits of change and therefore growth rates were decelerating. Longitudinal research over about 10 years might elucidate this important point.

The main implications of the cognitive and social findings reported here seem clear; we cannot predict a necessarily poor outcome for children with I.Q.'s in the 50's, 60's and 70's if they also come from very adverse environments. While there is no doubt much truth in Bowlby's thesis that separation from the mother may have very unfortunate effects on the child, we must not interpret this finding too rigidly for, as Hilda Lewis (1954) points out, it may be even more disastrous to leave some children with unsuitable parents. Moreover, findings such as these suggest that feebleminded persons, most of whom come from bad or very bad conditions, are within limits far from being the hopeless propositions which until recent years was generally accepted; already a much more positive attitude towards them is emerging.

(c) *Relevance of other work to the present findings*

That environment can exercise a stimulating or depressing effect upon intellectual function is now generally accepted (cf. Vernon, 1955), although there is still a dearth of knowledge about precise mechanisms, the limits to, and the factors limiting, such changes. So far as the mentally deficient are concerned, however, the concept of pseudo-feeblemindedness continues often to be used to explain away I.Q. increments among them. Perhaps the best statement of this view is that of Porteus (1941): 'Very wide differences in intellectual status merely indicate that the first diagnosis was wrong. Any child who finally functions at a normal level proves thereby that he never was feebleminded.' In effect, Doll's attitude (1941) is similar; essential incurability is suggested as a necessary criterion of mental defect; one might, therefore, have to wait 20 years to find out whether a child improved and hence no firm diagnosis of a high-grade case could be made in childhood. Both writers express an outlook firmly rooted upon the concept of a fairly rigid constancy of the I.Q., a view to which few would now subscribe. Elsewhere, (1955), we have argued that the concept of pseudo-feeblemindedness is now untenable in cases showing deferred maturation, since most studies show that initial testing gave no indication of underestimation at the time, and must be considered a true picture of the then present status of the person. Thus Porteus's view that *mistaken diagnosis* has occurred must be rejected; rather, the clinician has made a *mistaken prognosis*. And the present studies have shown that it is possible to predict I.Q. changes by an external criterion. It is important to realize, therefore, that arrest in intellectual development, or a decelerated growth rate, may be permanent or impermanent to varying degrees. In both types, the symptoms (e.g. social incompetence, behaviour disorder, or low I.Q.) may be similar, but it is a truism that the same symptom may arise from different causes.

Previous researches on I.Q. changes in the mentally deficient have been critically reviewed elsewhere (Clarke & Clarke, 1958). Here, it is proposed to discuss only one specific study relating to English feebleminded patients, and two others with broader implications.

(1) Mundy (1955), in studying environmental influences on the I.Q., used an experimental and control group each consisting of twenty-eight adult feebleminded women, and believed that the two groups were initially equated on a number of important variables. The interval between test and retest was 2 years, during which time the controls remained within a good institution (a convent) while the members of the experimental group enjoyed normal life experiences, mostly in residential wage-earning employment. On

retest, the control group showed an average increment on the Wechsler test of 2·25 points, and on the Progressive Matrices of 3·46 points. The average gains made by the experimental group, however, were 11 and 9·82 I.Q. points, respectively. Mundy believed that these results indicated that present environmental differences induced differential gains, but the validity of this conclusion clearly rests upon whether control and experimental groups were in fact *initially equal in all respects*. We are told, however, that there was no clinical selection by the hospital for outside employment, and that 'any able-bodied patient who wished for outside work was found suitable employment as soon as this was feasible'. This suggests that the members of the experimental group had all, through drive, ambition and initiative, asked for outside employment and hence gained it, while for several years at least the members of the control group had been content to remain within a good institution: there was thus a difference between the groups in terms of 'self-selection'. It should be added, however, that in a personal communication the author has stated that some cases were in effect originally transferred to employment for administrative reasons; this would reduce the amount of 'self-selection' but it might well be an indication of selection by the hospital. A further reason for believing that the control group contained atypical patients was the extraordinary stability of their test scores over the 2-year period; twenty-five out of the twenty-eight remained within 0 and plus 4 I.Q. points of their initial Wechsler result, and no patient increased by more than 6 points. This would imply a much higher test-retest correlation than is reported elsewhere in the literature. Because of these two factors, therefore, Mundy's conclusion is open to some doubt, and, in view of our own findings, it may well be that the I.Q. changes would have occurred differentially in any reasonable environment. Where her research and our own agree, however, is that early very adverse experiences played a part in retarding intellectual and general development. (See also Mundy, 1957.)

(2) Some striking parallels with our own findings emerge from experimental work in an entirely different field of deprivation, although these analogies ought not to be pressed too far. Widdowson & McCance (1954) studied the effects of different food diets on undernourished German orphanage children. Among other things, their growth in height, weight and ossification were investigated. The most striking observation was that following the period of undernourishment, these children grew even better than normal children of the same age, in spite of the very simple diet provided, which contained little milk or animal protein, and up to 75% of its calorie value in the form of flour products. It may be either that the flours of different extraction rates were equally nutritious or that the differences between them were masked by the exceptionally rapid growth rates of undernourished children when receiving a diet adequate in calories. Even the addition of a special milk supplement did not increase further these already accelerated rates of increase. In an appendix to this report (*op. cit.* p. 80) J. O. Irwin indicated that alternative viewpoints are equally possible. Because of original underfeeding, these children might have grown on any diet not grossly inadequate, the response to different amounts of vitamins would therefore be largely non-specific, and specific differences might still be detected in children with a more normal nutritional history on similar diets. On the other hand, others might be more struck with the fact that these children grew equally well on all the diets and at a faster rate than normally. The similarity between these concepts and those relating to our own study of environmental effects is very close. Another interesting finding, so far as weight increases were concerned, was that initial increases

were more rapid than later ones (cf. our *Second Study*). It is of some consequence that unusual growth rates (with respect to normals) after malnutrition should in some ways be analogous with unusual increments in cognitive and social development in persons who suffered a different form of deprivation.

(3) The work reviewed by Bowlby (1951) is too well known to need summary; it is highly relevant to our thesis that early adverse environmental influences have played a part in producing mental deficiency. His monograph is, however, almost entirely concerned with studies of children or young adolescents. In general, Bowlby is pessimistic about their recovery from early psychological damage, particularly in the absence of psychotherapy. Yet our own subjects, the majority of whom could scarcely have suffered worse childhood experiences (followed by institutional upbringing in many cases) being therefore *par excellence* a deprived group, have on the whole done much better than might have been expected, and none has received formal psychotherapy. Similarly, some of the social studies such as that of Charles (1953) have indicated that persons who were originally high-grade defectives in childhood do not develop according to a rigid stereotype, that a wide range of individual differences is maintained, with many eventually functioning in adult life at dull-normal or normal levels. There is no way of knowing what our subjects would have been like had they enjoyed normal conditions from birth onwards, but there is no doubt that in adult life they are, on the whole, very much less impaired than would have been expected from Bowlby's original theory.

Two very important general questions arise. First, to what extent are the main variables those of human susceptibility to early psychological damage and resilience thereafter, and what is their nature? Why is it that some children are utterly overwhelmed by adverse conditions while others are apparently stimulated by them? We have no real estimate of the proportion which emerges relatively unscathed, simply because we tend only to study the problem cases, but it may well be larger than we think. For example, often two or three children in a family come to the notice of the authorities while their other siblings do not, and we know next to nothing about the latter. Secondly, is it possible that, when more deprived children are followed up through adolescence and adult life, it will be found that, in many, the initial psychological damage, as with our own deprived population, tends steadily to be repaired to varying extents? This, if true, would represent a need for modification of the 'maternal deprivation' thesis, and it is of great interest that Bowlby, Ainsworth, Boston & Rosenbluth (1956) have already begun to re-formulate some aspects of it.

The writers are greatly indebted to Dr J. F. MacMahon, Physician Superintendent, for his many suggestions and constructive criticism; to Dr Gordon Claridge for having made independent assessments of social histories; and to Dr H. R. Beech for statistical advice and for the computations.

References

Bowlby, J. (1951). *Maternal Care and Mental Health.* Geneva: World Hlth Org.

Bowlby, J., Ainsworth, M., Boston, M. & Rosenbluth, D. (1956). The effects of mother-child separation: a follow-up study. *Brit. J. Med. Psychol.* **29**, 211–47.

Charles, D. C. (1953). Ability and accomplishment of persons earlier judged mentally deficient. *Genet. Psychol. Monogr.* **47**, 3–71.

Clarke, A. D. B. & Clarke, A. M. (1953). How constant is the I.Q.? *Lancet*, ii, 877–80.

Clarke, A. D. B. & Clarke, A. M. (1954*a*). A rehabilitation programme for the certified mental defective. *Ment. Hlth, Lond.*, **14**, 4–10.

Clarke, A. D. B. & Clarke, A. M. (1954*b*). Cognitive changes in the feebleminded. *Brit. J. Psychol.* **45**, 173–9.

Clarke, A. D. B. & Clarke, A. M. (1955). Pseudo-feeblemindedness—some implications. *Amer. J. Ment. Defic.* **59**, 507–9.

Clarke, A. M. & Clarke, A. D. B. (Ed.) (1958). *Mental Deficiency: the Changing Outlook.* London: Methuen. (In the Press.)

Doll, E. A. (1941). The essentials of an inclusive concept of mental deficiency. *Amer. J. Ment. Defic.* **46**, 214–19.

Eysenck, H. J. (1953). *Uses and Abuses of Psychology.* London: Penguin.

Hebb, D. O. (1949). *The Organisation of Behavior.* London: Chapman and Hall.

Lewis, H. (1954). *Deprived Children.* London: Oxford University Press.

McNemar, Q. (1940). A critical examination of the University of Iowa studies of environmental influences upon the I.Q. *Psychol. Bull.* **37**, 63–92.

Mundy, L. (1955). Environmental influences in intellectual function as measured by intelligence tests. London University, unpublished M.Sc. Thesis.

Mundy, L. (1957). Environmental influence on intellectual function as measured by intelligence tests. *Brit. J. med. Psychol.* **30**, 194–201.

Nemzek, C. L. (1933). The constancy of the I.Q. *Psychol. Bull.* **30**, 143–68.

Penrose, L. S. (1949). *The Biology of Mental Defect.* London: Sidgwick and Jackson.

Porteus, S. D. (1941). *The Practice of Clinical Psychology.* New York: American Book Co.

Thorndike, R. L. (1933). The effect of the interval between test and retest on the constancy of the I.Q. *J. Educ. Psychol.* **24**, 543–9.

Thorndike, R. L. (1940). 'Constancy' of the I.Q. *Psychol. Bull.* **37**, 167–86.

Vernon, P. E. (1954). Symposium on the effects of coaching and practice in intelligence tests. V. Conclusions. *Brit. J. Educ. Psychol.* **24**, 57–63.

Vernon, P. E. (1955). The assessment of children. In *Studies in Education*, no. 7. University of London Institute of Education.

Widdowson, E. M. & McCance, R. A. (1954). Studies on the nutritive value of bread and on the effect of variations in the extraction rate of flour on undernourished children. *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 287. London: H.M.S.O.

(*Manuscript received* 14 *March* 1957)

SHAPE RECOGNITION: A REPLY TO DEUTSCH

By P. C. DODWELL

Psychology Department, Birkbeck College, London

The main points made by Deutsch (1958) call for some comment. First, there is ample evidence in Lashley's (1938) reports to show that his rats were *in fact* discriminating by a part of the visual display in most cases and not by the whole; this would seem to indicate that the rats were not basing discriminations on the properties of complete shapes, whether those shapes could be 'seen whole' or not. Admittedly, putting the argument in terms of the rat's (inferred) field of distinct vision is weak, but this does not affect the fact that Lashley demonstrated a small field of *effective* vision in his experiments. Secondly, there is some evidence to show that rats do not treat rotated shapes as equivalent, although they may learn to do so. Fields (1932), in his investigation of the rat's ability to discriminate a triangle in 24 different positions (rotations) from a circle, found that the average number of trials needed for the initial discrimination was 580; for the first rotated position, 230; for the next 11 positions, an average of 45 trials each; and for the remaining 11 positions, 21 trials each. This suggests that the rat may be learning to treat a family of related signals in the final common cable (such as would be generated in the writer's system) as equivalent to one another. It is difficult to see why such learning should be necessary on Deutsch's theory, since the messages for all positions of rotation would be the same. Furthermore, on pairing a square or rectangle (base horizontal) with a rotated triangle, the animals tended to choose the shape with horizontal base in preference to the triangle. In the writer's theory such a 'wrong' choice is explicable, since the signal for the square or rectangle has one feature in common with the signal for the first position of the triangle—a long delay on the horizontal f.c.c. attributable to a horizontal base. In Deutsch's theory such a choice could not be explained. This evidence is not conclusive, however, since Fields did not demonstrate whether his rats were responding to parts of the display or the whole as Lashley did. In Fields's experiment, the visual angle subtended was approximately 18°.

Deutsch says that there is evidence that rats jump spontaneously to horizontal bars, but does not quote an experiment which shows this. It was not the case in the writer's experiment, although the horizontal/vertical discrimination was quickly learned. Deutsch suggested a plausible reason why Lashley's rats could discriminate a square rotated through 45° from a circle but not a square from a circle, and this same argument *could* explain why they discriminated a square from another rotated through 45°. Indeed, if Deutsch's theory holds, such an explanation (or something similar) would be needed. Gellermann's (1933) finding is open to the same interpretation as that given above for Fields: the signal for the rotated triangle is 'recognized' as belonging to the same family as the signal for a triangle with base horizontal, and the receptors are rotated to confirm this. In Deutsch's theory it is difficult to see what purpose such a rotation would serve.

It is true that Deutsch can incorporate a system for separate coding of information about tilt in the manner he suggests; however, the evidence seems to point to a coding

system (at least in visually primitive organisms) in which information about shape and tilt are compounded in one signal, or give rise to signals which mutually interfere with each other, which is not the case in Deutsch's (modified) theory.

REFERENCES

DEUTSCH, J. A. (1958). Shape recognition: a reply to Dodwell. *Brit. J. Psychol.* **49**, 70–1.

FIELDS, P. E. (1932). Studies in concept formation. I. The development of the concept of triangularity by the white rat. *Comp. Psychol. Monogr.* **9**, 1–70.

GELLERMANN, L. W. (1933). Form discrimination in chimpanzees and two-year-old children. *J. Genet. Psychol.* **42**, 3–27.

LASHLEY, K. S. (1938). The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.* **43**, 123–93.

(*Manuscript received* 4 *May* 1957)

CRITICAL NOTICE

By JEROME S. BRUNER
Harvard University

Thinking: An Experimental and Social Study. By Sir Frederic Bartlett. London: George Allen and Unwin. 1958. Pp. 203. 18*s*.

Sir Frederic Bartlett is a man remarkable for the combination of common sense and high adventure that has always informed his work. The interplay of shrewd conservatism and speculative daring in this, his most important work since *Remembering*, reminds one of nothing so much as the blend of loyalism and freebooting in those earlier Elizabethans, the privateers. He captures prizes, but he loses not his head or his calm. This is a remarkable book, 'vintage Bartlett' as Sir Frederic's students would say, full of insights and ambiguities. If it does not succeed in taming its subject, *thinking*, it certainly leaves it well worked over. A reviewer may perhaps be forgiven for introducing a book in so biographical a view; for this is in such a deep sense a biographical book. Not only is the reader witness to a dazzling display of uniquely Bartlettian intellectual technique, but the author offers among others his own past efforts (the construction of *Remembering*), as laboratory material on which to try out his notions about the process of 'adventurous thinking'.

The point of view, the theoretical line of the book finds its origin in the analogy between skilled bodily performance and the process of thinking. Sir Frederic does not insist on a parallel between the two, but rather sees the latter as an extension of the former, one made possible by the capacity of man to manipulate symbols rather than being bound to the employment of receptor-effector acts. Both bodily skill and effective thinking depend upon a fitting use of evidence, the former requiring sensory regulation, the latter something more difficult to define. Each is marked by a sequential integration of component processes, and in consequence, each can be shown to exhibit certain common or analogous properties. One of these is *timing* in the use of information and in the executing of steps, be they muscular acts or acts of inference. Each type of sequence, moreover, has its *stationary phases*, its critical 'hoverings' and 'halts' during which timing is regulated. Betimes, each sequence reaches a 'point of no return', the point at which the batsman can no longer change his stroke without the price of disruption, the point at which the thinker becomes the victim of what he has done to the evidence with which he has had commerce. Throughout, each process is pregnant with its terminus; moving toward an issue—be it the drawing of a conclusion or the delivery of a drop shot. This is the '*direction*' of the sequence, and upon it Prof. Bartlett rests much of the theoretical edifice he constructs.

Since, I believe, the book gets into some difficulties by relying on the parallel between bodily skill and thinking, a word needs saying about the difference between the two, and how it is that Prof. Bartlett conceives the one developing from the other. One finds the clue to the difference in a close examination of the nature of 'direction'. In its simplest form, direction is identified with 'receptor-effector anticipation' in which perceptual scanning runs a bit ahead of the action to be performed with the result that reaction has a seemingly preternatural swiftness. If, for example, we are to move toggle switches in the direction indicated by series of independent and discrete displays, the time required for the making of a movement is absurdly little less than the time it takes the operator merely to observe or interpret what the display is saying. But this is a narrow range of anticipation, very small beer. Bartlett offers a second stage en route to thinking, 'perceptual anticipation', in which the operator deals not with single and independent events but with the directional properties of sequences—with sequential redundancy, if you will. Such sequential anticipation is possible *provided* the operator has created a model or schema to represent the recurrent features of the sequence, a problem of learning and information storage toward which Prof. Bartlett is oddly cavalier. The author suggests that 'perceptual anticipation comes into the picture because the behaviourally more basic receptor-effector anticipation very often breaks down'. If this is true, then it should also be true that schematizing in memory occurs because efforts to deal with single items by rote 'very often break down'—a view to which the author of *Remembering* should take violent exception, even if it is stated by the author of *Thinking*!

Finally, 'when we discover or learn how to use warning signals or symbols' in dealing with the evidence

provided by sequences of events, then we are dealing with the kind of direction that is most characteristic of thinking.

What, then, is thinking? 'The process begins when evidence or information is available which is treated as possessing gaps, or as being incomplete. The gaps are then filled up, or that part of the information which is incomplete is completed. This is done by an extension or supplementation of the evidence, which remains in accord with the evidence (or claims to do so), but carries it further by utilizing other sources of information besides those which started the whole process going, and, in many instances, in addition to those which can be directly identified in the external surroundings. Between the initial information and the terminal stage, when the gaps are alleged to be filled, or completeness achieved, theoretically there is always a succession of interconnected steps...they, more than anything else, are what makes an experimental approach to thinking theoretically possible, and they confer upon thinking its character of necessity' (p. 75).

Now, like most working definitions, this one is not a particularly felicitous one. For as C. S. Peirce so wisely insisted long ago, and Wittgenstein more recently, the definition of a thing or process is the array of predicates one can launch in the sentence frame 'if x, then...'. But it is adequate to underline what Prof. Bartlett conceives to be formally central: that there be evidence with gaps and that these gaps be filled in a manner which, to the thinker, is 'fitting' or 'adequate'. The gaps exist, we may infer, by virtue of 'direction'. Without direction there would be such an entropic hodge-podge that one could not tell the gap from the chaos. Which leads me to my first major criticism of this provocative work. It is that too often (not always), the author adopts a kind of Baconian or common-sense view to the effect that direction (or what I would prefer to call 'structure') somehow grows out of evidence as the autochthonous gods were supposed to grow out of the hallowed soil of Olympus. At his most negligent he speaks merely of the 'extraction of rules' from evidence, without examining in sufficient detail what is involved in such extraction.

What are the great principles of transforming informational input, taming it into a structure, imposing direction upon it? I believe that the 'skill' paradigm to which Sir Frederic has wedded this enterprise keeps him from staring boldly enough at this problem. When he returns to the familiar hunting ground of memory organization, the sparks fly. 'All the way through this book I have taken the view that thinking is possible only if much that precedes it in an order of development continues to exercise an influence. This is true in that the thinker is able to avail himself of evidence no longer a part of his immediate external environment, and in that the ways in which he treats the evidence very often have conventional characters which could not possibly be constructed on the spur of the moment. Perhaps the very most critical of thinking problems have to do with what sort of processes of storage of evidence thinking demands, and what are the implications of the uses thinking can make of the stored evidence' (p. 87). Here would have been the staging area for a revisit to the *schemata* which stood Sir Frederic in such good stead in his classic volume twenty-five years ago. As it is, he has some most arresting things to say about the grasping of similarities in contrast to differences for the construction of viable thought structures (a term he does not use), viable meaning 'transferable'. See, for example, one specific difficulty introduced by the 'skill model'. Sir Frederic notes that, given direction in thinking, there is a point of no return as in skilled movement. This is indeed the case. But it is also the case that the models that we use in locating direction have as one of their most interesting properties the character of reversibility. Two plus two equals four, but four is also decomposable into these two factors. 'Some p's are q' has the reverse implication that 'some or all q's are p'. The structures that underlie thinking have neither a property of timing nor of 'no return'. It is only when we base an hypothesis upon them that these properties emerge. As I shall urge again before ending this review, the properties of these underlying structures or schemata are much in need of study if we are to understand the thinking acts that are based upon them.

The book deals with four types of thinking, each of which may be conceived to have a somewhat different objective and each a recognizable strategy. The first is 'closed-system thinking'. In rough terms, closed-system thinking involves the attempted solution of well-defined problems of a formal or highly structured nature, and such attempted solutions involve either the interpolation of information in a structure, as in supplying the numbers between 2, 4, 8 and 128, 256, 512 or in extrapolating from a series such as 1, 3, 5, 7, 9, 11, 13,.... Closed-system thinking may also involve transforming 'evidence in disguise', which consists of recognizing the relevance of information to a problem although the information is organized in a fashion not immediately relevant to the problem. Or it may mean that the information must be manipulated in order for a problem to be recognized.

Prof. Bartlett's demonstrational experiments in this portion of the book, fresh and ingenious though

they be, leave one a bit puzzled. They show beyond question that interpolation and extrapolation occur, and that subjects use a highly limited array of solutions to such problems. Why? Surely it is not for Kantian reasons, that there are only a limited number of categories of thinking. I could not help but ask myself what it was that we knew about the *acquisition* of the solution models that people apply to closed-system problems. Very little indeed. Perhaps the answers must come from studies of the learning or acquisition of 'rules' or 'logical models'. One misses reference to such work, particularly the researches of Piaget.

A second type of thinking is more 'adventurous' and is called *experimental*. 'His [the thinker's] aim is not simply to extend and define knowledge within a system that is to be treated as already completed, but to help to develop and understand systems which are so far literally incomplete' (p. 98). The discussion of experimental thinking is, in the reviewer's opinion, the wisest and most penetrating set of reflections on the subject that has ever been put to print –flexible, informed by much direct experience as well as by some interesting experiment, full of challenging notions about the use of past experience in transfer. Much emphasis is placed on strategies for spotting the similarities between events as a basis for discovering common features in seemingly disparate problems. For it is the discovery of such common features that makes transfer of solution possible. Here I think Sir Frederic has pointed his finger directly at the reason why in thinking and conceptualizing, people prefer to work with 'positive instances', examples that illustrate the principle they are working on. For these permit us to transfer what we have already learned, whereas negative instances leave us with nothing to transfer, force us back into a kind of trial and error. As James Bryant Conant once remarked, the object of science is to reduce empiricism—to allow us to spot and exploit the similarity between that which we encounter anew and what we already know.

The third type of thinking discussed, 'everyday thinking', has as its aim something else, something of central import to the social psychologist. 'The closed-system thinker and the experimental scientist are set to achieve results which carry with them an internal compulsion. Everyday thinking also aims to arrive at a compelling issue, and, however it makes its journey, to stop where every other normal person must stop. But the steps to the resting place are never all of them exhibited, and there is no inherent compulsion in their nature and relations. It is not in the nature of the assertion, but in the manner of assertiveness that everyday thinking seeks to attain necessity...the source of the compulsion being now within the thinker, and particularly in his social group, it is possible, and indeed common, for completely contradictory issues to claim the same necessity.' Sir Frederic then goes on to say that everyday thinking might as appropriately be called 'immediate communication thinking'. 'Immediate, here, means dealing with highly topical events....The communications are then likely to share the urgency of the events, and thus to favour a strong and unconditional mode of expression. The same mode is still further strengthened by all those self-regarding sentiments from which nobody who enters into the "immediate communication" situation can escape.' In achieving generalization, even in the absence of social convention, it proceeds not by a 'painstaking, slow, and intelligent comparison of many instances, and the abstraction of common features, but by an easy and natural exaggeration which in the extreme may be based upon one single impressive personal experience...where formal or experimental structures are concerned..., the few instances are examined for whatever "rules" they may reveal. In everyday thinking the instance, or the few instances, are not so much explored as, so to speak, "savored" and used. The generalizations of the non-exploratory kind have little to do with transfer of practice or training save that they make it more difficult.'

The discussion and observations on everyday thinking are highly challenging. My only complaint is that Prof. Bartlett makes social communication seem like such a *barrier* to adequate thinking. As Arieti has recently argued (and Piaget, Vigotsky and George Herbert Mead before), it is the need for adequate social communication that leads people to learn and follow the rules of logic and inference as well as conforming to the primitive principles of denotation. Finally, Prof. Bartlett would have found a powerful ally for his contentions about the conventionalizing effects of social communication in Benjamin Lee Whorf whose work on psycholinguistics shows how deeply the socialized forms of assertiveness affect the texture of thinking.

Let me finally revert to several general points that are put in issue by this challenging and pioneering work. The first has to do with the manner in which the process of thinking itself may be most fruitfully formulated. As Sir Frederic insists, thinking involves the use of past experience, and I would add *organized* past experience. We organize our past experience in such a way as to get 'maximum travel' from it. Such organization results in our being able to subsume new experience under old rubrics so that we need not deal with the new experience *de novo*. We convert past learning into something generic,

and having grasped that $A > B$ and $B > C$ implies that $A > C$, it makes no difference whether we are dealing with ergs, potatoes, or Irish whisky. To be sure, the structures or schemata used in thinking are altered by the range of instances to which they are applied, but what is basic to thinking is that such underlying structures or 'rules' or 'transforms' are necessary.

The process of thinking, then, may be conceived in two steps: the one is a process of rule or structure learning, the second is a process whereby we recognize events as conforming to the structures we have learned. It is about the former process that Prof. Bartlett writes in discussing 'adventurous thinking', about the latter in connexion with closed-system thinking. Yet, one gets the impression that there is some confusion between the two that leads him to speak of extracting 'rules from evidence' in discussing the way subjects deal, for example, with well-defined arithmetic problems.

I would urge that at least two missions should be marked out for the psychologist concerned with thinking. One is to learn how organisms acquire—invent, discover, learn, mature into—the structures or schemata that we apply in thinking and problem solving, The genetic approach is one fruitful avenue, most recently attested to by *De la logique de l'enfant à la logique de l'adolescent* of Inhelder and Piaget. A second approach is phylogenetic, and I would cite Klüver's *Behavior Mechanisms in Monkeys* as a stunning example. Prof. Bartlett's account of adventurous or experimental thinking is yet another approach—a close analysis of how people operate in their thinking when they are at the frontiers of knowledge. His exciting account of the development of *Remembering* is a prime document.

The second mission is to examine the strategies whereby people are enabled to recognize events as exemplifying the principles or structures they have learned so that these principles or structures may be applied. How, in short, do we achieve what William James once called 'the electric sense of analogy'? How do we avoid seduction by the irrelevant, confusion by the seemingly disordered, etc? Sir Frederic's chapters on thinking in closed-systems suggest some interesting answers to these questions.

In fine, *Thinking* is a book whose appearance gives promise that the experimental psychology of thinking is emerging as a vigorous field of inquiry. It is a bold overture, one that will have an impact on future research.

PUBLICATIONS RECENTLY RECEIVED

Schools of Psychoanalytic Thought. By RUTH L. MUNROE. *London: Hutchinson.* 1957. Pp. xvi+670. 42*s.*

The authoress of this book is a clinical psychologist in New York, specializing in diagnostic tests. She tells us that she has collaborated professionally with 'some fifty' psychoanalysts 'of various schools'; that she is the wife of Dr Bela Mittlemann, an established analyst, with the marriage perquisite of being able to listen-in to shop talk; and that 'as part of my training as clinical psychologist as well as for personal reasons, I have been in analysis with a fairly classical Freudian and at another period with one of Horney's first lieutenants'. Her chief purpose in writing this book seems to have been to evaluate and compare those analysts 'whom most of us in America think of as having established schools'. To do this, she has carefully and strenuously attempted to present 'an objective account' of these different schools. 'In writing', she tells us, 'I have tried to feel myself into each system as a whole, thinking for the nonce in the terms of its major premises and working through to its conclusions as if I were successively a Freudian, an Adlerian...' and so on. In her attempt to achieve expository accuracy and fairmindedness, she took the step of submitting her draft account of each school 'to experts selected for their special acquaintance with the school'. When she comes to evaluate and compare, she adopts 'a critical stance' and attempts an 'integration' of the schools she has presented. In doing this, she passes from exposition to clearly separated comment. This comment, she says, is of 'two orders, often closely commingled'. First, she presents the reflexions of a clinical psychologist about a school in order 'to invite thought and inquiry in areas which, I feel, have been neglected by the school'. And secondly, she attempts 'a formalistic rephrasing of the basic problems of psychological science' as a means of 'integrating the schools reported' in this book. If I understand her rightly, this rephrasing is part of an attempt to develop a more inclusive theory or perspective than the theories of the schools that she reports. For this purpose, she has 'borrowed many ideas from contemporary philosophers, notably from the romanticists and the logical positivists'.

The plan of the book is as follows. She begins with a short introduction in which she explains the general nature of the enterprise. In the next chapter, she extracts and explains the concepts that she thinks basic to and accepted by *all* brands of psychoanalysis. These are four in number: psychological determinism, the role of the Unconscious, the goal-directed character of human behaviour, and the genetic approach to personality. She then goes on to consider the school of 'Freud and the "Freudians"'; and this makes up the largest single section of the book. By 'Freudians' she means those analysts who accept the 'libido theory'. She then examines the 'non-libido schools' of Adler, Horney, Fromm and Sullivan. Her main concern is with the cleavage in the United States between 'libidinal' and 'non-libidinal' analysts. She continues with a short examination of Jung and Rank, who 'do not fit this libido-non-libido dichotomy'. She ends with an 'Epilogue' in which she puts together her suggestions for integrating the psychoanalytic theories she has examined.

I shall concentrate on the chief concern of the book. What are the main comments she makes on the libido and non-libido schools? What integrating 'rephrasing' does she offer? I pick out three points and themes on which she criticizes Freud and the 'Freudians'. (*a*) This theory, even in its current development, overlooks or pays insufficient attention to 'drive systems' that are non-sexual in character. By a 'drive system', Munroe appears to refer to a systematic pattern of goal-directed activity, closely connected with some bodily apparatus whose functioning 'involves inner tensions', and which are 'oriented toward the external world' for the completion of the systematic pattern of activity. Obvious examples are the respiratory, cardiac and digestive systems of the child; less obvious examples, perhaps, are its neuro-muscular (or 'motility') and speech systems. In overlooking these other non-sexual systems, Freudians are apt to forget (in theory if not in practice) that 'from the very beginning the infant is dandled, talked to and stimulated visually, usually by the adults who also care for its "sexual" systems'. Because of this, Freudian theory misses much in the relationship between child and adult that may be, and probably is, important. Munroe suggests, for example, that a good relationship between them and, consequently, a favourable development in the non-sexual systems 'will suffice in the majority of cases of libidinal development'. In its exclusive emphasis on the erogenous zones, Freudian theory is guilty

of a major error in Munroe's eyes. This is the error of 'reductionism'. To commit it is 'to apply a perfectly sound idea too generally, with neglect or exclusion of concomitantly operating systems of events'. Scientific theorizing, she argues, should be multi-dimensional in character. It should study data 'along many lines (dimensions), which are then related to one another'; and thereby avoid the type of integration that 'reduces "everything" to a master principle'. (*b*) Early Freudian theory is incomplete in not giving enough recognition to the importance of 'ego' functioning. She discusses Hartmann's suggestions with approval. She maintains that there is 'a major omission from Freud's statement of the important dynamic groupings of systems (structures, institutions) that come to function more or less independently in the mature personality. This is the self image', and she explains the relationship between this concept and the super-ego and the (neglected) ego-ideal of Freudian theory. (*c*) In discussing pathology, she declares that 'for every case that neatly fits the simple Freudian dynamics here outlined, there are ten for which the interpretation requires drastic modification....Any simple extrapolation from libido theory to the majority of cases encountered in unselected practice comes to grief very quickly.'

Her comments on the non-libido schools complement those she makes on the Freudian. Consider Horney alone. She approves of Horney's concept of the 'idealized image of the self'. This and what goes with it encourages the development of 'ego-psychology', with its emphasis on the interpersonal relations of the here and now; and she remarks that Freudians 'are just beginning to give proper theoretical weight to concepts which Horney has been shouting from the house tops for twenty years'. But she criticizes Horney on two counts among others. (*a*) She commits the error of reductionism in making 'the need for security' the basic principle by reference to which all the variations in personality development are elucidated. (*b*) She is quite wrong in rejecting the Freudian thesis that the sexual drive systems have an important role *per se* in personality development and later dynamics. (The same criticism applies, *mutatis mutandis*, to the other theorists of the libidinal school.) 'There is every reason', Munroe writes, 'to suppose that the biologically determined sexual systems *do* have an autonomous course of development which must profoundly affect the relationship of the young child to his environment during the early months and years of his life—and therefore profoundly affect the attitudes he develops along the lines emphasized by Horney herself.' In view of Munroe's own emphasis on the importance of the non-sexual systems, she goes out of her way to declare that ' I cannot state too emphatically my judgment that the infant *is* overwhelmingly concerned with its profound biological functions (oral, anal and genital); that its early interpretation of the world *must* be primarily constellated round these functions'. This dual attitude enables her to say that 'there are probably relatively few cases (i.e. children) in which the sexual systems *as such* actually provide the overwhelming focus for the personality as a whole in the manner suggested by the earlier Freudians'; *but* that 'there are also relatively few cases in which the developmental systems that these schools emphasize so thoroughly determine the personality as to make the course of the sexual systems in themselves a matter of indifference'.

These comments by Munroe help us to understand the sort of 'integration' she espouses. She urges on us a more inclusive 'perspective' in which emphasis is laid on the organismic self (or total personality) operating through different sorts of 'systems'—the 'drive systems' of the organism, the 'dynamic sub-systems of very great variety' which develop in the course of living, and the 'transient sub-systems' established moment by moment. This organismic self develops 'a self image' which has 'an underlying dynamics' of its own in our culture that merits study. Our drive systems are sexual and non-sexual in character, and the latter also have 'an autonomous dynamics' that requires further investigation. What is essential is that we avoid premature universalization and tight, rationalistic theorizing. Her more inclusive perspective and rephrasing can be best appreciated, perhaps, by noting how she rejects the one-sidedness of both the Freudian and non-libido schools, and how she attempts to synthesize their constructive contributions. She ends by extending her 'view of systems' to include the difficult problem of aggression—on which she has interesting suggestions to offer.

My brief sketch of her argument does not begin to do justice to the mass of comment and suggestion the book contains, or to the 'integrative rephrasing' she offers. I have said nothing, for instance, about her approval of Kardiner and Anna Freud, of her criticisms of Melanie Klein, and of her remarks on Sullivan (whose line she considers the most promising of the non-libido group), and of her recognition that Jung's *persona* is close to her concept of self-image. In this review I have only been able to pick on one central contribution that the book offers us. It is fair, I think, to warn the reader that, in her endeavours to be fairminded, Munroe has been longwinded and a bit repetitious—the book has the laboured air of careful toil at midnight.

What are we to make of the central contribution I have outlined? Let us ask the preliminary question: 'Whom is she addressing?' In so far as she is writing for young psychiatric registrars, for students of

psychiatry and clinical psychology, for the younger psychoanalysts, and the like, she will probably not be unsuccessful in inducing them to be more tolerant of alternative points of view in psychoanalysis, and to appreciate better the contributions that each has to make. Hence the book may do a little to bring about a rapprochement between the practitioners of the various schools. But in so far as she is writing for established analysts, for people who have been through years of training (or 'indoctrination', as some allege) in some School or Institute of psychoanalysis, I doubt whether the impact of the book will be great. Likewise, in so far as she is addressing the hardheaded, empirically minded layman, or scientific worker, I doubt whether the impact will be at all considerable.

I do not think the book will move the established analyst for the following reason. When I ask myself: 'What grounds does she offer for her integrative line?', I find it difficult to discover any. I suspect that the grounds that move Munroe to talk as she does are: (1) that this line synthesizes her own clinical experience and the alternative psychoanalytic theories in a way she finds personally satisfying; (2) a hunch that her line will help the development of psychoanalysis at this juncture. If I am right about this, then it is evident that the whole book has the strength and weakness of the usual story produced by a clinical psychologist or psychiatrist about a patient or a set of clinical data. The story may be very suggestive and illuminating, and it may even cover all the known facts coherently. But it is not very compelling, especially to established practitioners with developed and pet stories of their own that they prefer to spin. What *would*, I think, move them would be to show them that some important thesis of theirs is *false*. I take it that she wants to say, for example, that the Freudian thesis about the exclusive role of the sexual drive systems is a false thesis. But does she *show* this? I hesitate to say so. What is rather futile is for Munroe to accuse the Freudians or Horney, or anyone else of 'reductionism' when she does not provide an independent criterion to decide whether an analyst is 'reducing' or not. In the absence of such a criterion, the word 'reductionism' functions, largely though not wholly, as Munroe's favourite smear word.

The book is not likely to move the empirically hardboiled for similar reasons. They are only too likely to ask: 'Why all these different schools of psychoanalysis? Why all the disputation? Especially as Munroe speaks about the techniques of the psychoanalyst as if they were quite satisfactory?' It is not enough to answer (as she might): 'Because psychological science must work with loose theories or systems, unlike physics.' For what, precisely, makes a system a 'loose' one in science? It looks as if the theories, or systems, of psychoanalysts are *so* loose that it is not possible in principle to take any steps to decide between them, in which case they cease to be loose systems *in science* and become systems of a literary and metaphysical type. But she would repudiate this contention. She wants to talk about a psychoanalytic theory 'as a provisional systematisation of events', rather like one of the theories of light, which means that it makes some claim to objective adequacy. But then she has to meet the challenge of verification—how does one set about confirming or disconfirming any one of these theories? She appears by implication to hold that this can be done by means of the data of the analytic session and the projective test, and the data from the direct observation of children. She makes constant and (to me) inconclusive references to data of this sort. Thus, she seems to imply that this sort of data can establish that motility, and other non-sexual systems, play the role she alleges they do; and that the infant's early interpretation of the world is constellated round its oral, anal and genital functions. But to take this line is to plunge into all the complexities of psychoanalytic method and all the uncertainties involved in interpreting the behaviour of children. Munroe does not deal with these complexities and uncertainties. So the book is likely to strike the hardheaded as theoretically and empirically shoddy. I think the book raises very difficult issues about the whole character of psychoanalysis and psychoanalytic theories. These are issues of logic, and I cannot help feeling that they are beyond the range of Munroe's competence.

But I have carped enough. I end with a note of congratulation. In writing this book Munroe tried to think herself into the various theoretical positions in psychoanalysis. I think she deserves to be congratulated on this aspect of her enterprise, and on the success she has achieved. To think oneself into the various theories in some domain is not an easy task, especially where some of the theories are personally repugnant. But it is an excellent discipline. I recommend it to any psychoanalyst and psychologist of personality, and especially to those of the combative type.

B. A. FARRELL

Backwardness in Reading. By M. D. VERNON. London: Cambridge University Press. 1957. Pp. viii + 228. 25*s*.

In her latest book, Prof. M. D. Vernon has produced a scholarly, fully documented and impartial appraisal of the present-day state of our knowledge about children's backwardness in reading. Although she deals with many technical researches, she has set out the material in a clear way so that the book

is not only of help to educational psychologists, but also to classroom teachers. The latter may lack some of the knowledge required in connexion with some of the more physiological and psychological aspects of the problem, but will, none the less, be greatly helped by Dr Vernon's clear statements and her full discussion of the educational issues involved.

Dr Vernon takes us first through research on the perceptual basis of reading and these chapters could well be read by psychologists for useful evidence relating to perception. She then discusses the nature of reading disability, emphasizing the fundamental difficulty experienced by backward readers in connecting up visual perception with auditory perception and with the cognitive generalizations and meanings it is hoped to develop. In discussing the innate factors underlying disability in reading, it is heartening to find Dr Vernon writing so firmly about the limitations of the blanket term 'congenital word blindness'. The term really conceals ignorance and the less it is resorted to, the healthier our research into the problem of reading difficulty will be. She then takes up the effect of various acquired difficulties in reading ability. This is an informative chapter and one can question her views only in one small respect: this is on page 130 where she hazards a guess at the effect of premature birth upon linguistic and visual functions. A recent large-scale survey in Birmingham of 7 to 8-year olds all born prematurely, *excluding* all cases of brain damage and other abnormality, revealed no significant difference in scholastic attainment between the premature sample and the control. In the last two chapters, Dr Vernon discusses environmental factors springing from the home and school and also the problem of the remedial treatment of reading disability. In the case of the problem of remedial education, she presents the material in a detached and well-balanced way. She wisely eschews a discussion of the more controversial matter of the best measure of backwardness.

Dr Vernon provides the reader with a critical conclusion at the end of each section and then sums up in a last section on the over-all position, and she squarely faces up to the fact that methods of experimental and clinical study have now 'failed to isolate any factors which appear universally in all cases of reading backwardness—other than inability to read easily!' Later we are told that 'the only fact which appears clearly at first sight is the heterogeneity of cases of reading disability—heterogeneous both in the origin and in the nature of their disability'. Causation is multiple and perhaps the only general conclusion to be made is that reading demands intellectual capacity and that disability in reading is shown by 'the child's general state of doubt and confusion as to the relationship between the printed shapes of words, their sounds and their meanings'. Dr Vernon also emphasizes the highly important influence of motivation in reading attainment, pointing out that maladjustment and reading failure go together, but that the evidence is insufficient to show how far one is caused by the other. This book is to be recommended as a comprehensive assessment of our present knowledge about reading disability. Prof. Vernon has provided us with yet another scholarly analysis of an important field in psychology and education.

E. A. PEEL

The Slow Learner: Some Educational Principles and Policies. By M. F. Cleugh. London: Methuen. 1957. Pp. vii+186. 13*s.* 6*d.*

The subtitle 'some educational principles and policies' is less misleading about the contents than the main title. The book is not about method, and is especially intended for administrators, who, according to the blurb, 'have the responsibility of organizing the provisions for special educational treatment'. As a result, Miss Cleugh ignores the basic problem of organization, which hinges upon the responsibility of educational psychologists in these areas. Even when a plan for systematic ascertainment is reported and commended by the author, the fact that it was carried out by a psychologist is mentioned only incidentally. Similarly, the book contains many examples of misclassification by medical officers and yet the obvious first rectifying step of making psychologists at least informally responsible for intelligence testing is not suggested.

Miss Cleugh's intentions are admirable and she successfully simplifies some of the complexities, while emphasizing the varied possibilities of helping dull and retarded children. The book can therefore be recommended as an introductory text for those students contemplating a career in educational psychology, although some of its language and ideas are rather out of date. In a discussion of I.Q.'s, children are described as having 'half the normal intelligence', and this may well encourage incorrect conceptions of test results. Miss Cleugh also takes it for granted that all mongols are ineducable, whereas most practising psychologists will have met exceptions to this rule. When it comes to emotional reasons for retardation, the author subscribes to the idea that it is pointless to deal with a symptom while neglecting

a cause. This psychiatric creed is open to experiment and the results might have important effects on the future development of child guidance. In the opinion of many teachers the best thing to do with children retarded in reading is to teach them reading, even though the original cause of retardation was quarrelling parents, a traumatic experience or what-have-you. In fact, one might go further and say that remedying the result might affect the cause, for it would not be surprising to find parents quarrelling less after their children had shown an unexpected academic success.

DAVID SCHONFIELD

The Dynamics of Anxiety and Hysteria: an Experimental Application of Modern Learning Theory to Psychiatry. By H. J. EYSENCK. London: Routledge and Kegan Paul. Pp. xiv+311. 1957. 32*s*.

Two very interesting trends in the recent history of psychology are illustrated by this book. The first is the growing feeling that it is high time for a *rapprochement* between the sort of psychology that looks for general laws governing all human (or mammalian or animal) behaviour and the sort that studies individual differences. Hitherto they have been pursued by distinct groups of psychologists, who have not kept abreast of each other's work and who have used such vastly different languages and methods that one would scarcely believe both lines of research to be directed at the same organisms. Eysenck intends the present work as a contribution to the removal of this anomaly. It represents, he says, a transition from 'classification to a study of dynamics...from nosology to aetiology, from description to causation', as far as his own development is concerned.

The second trend is the spread of a changed view of the relations between learning theory and psychoanalysis. At first, learning theorists were impressed with the common ground between their conclusions and those of Freud, despite the very different kinds of data from which the two were derived. They produced theories of human behaviour, methods of psychotherapy and accounts of what happens in psychotherapy which, while they embodied many divergences from Freud's matter and especially from his phraseology, preserved much of the picture that he had painted. The group associated with Hull at Yale were particularly diligent in this regard, but Tolman, Skinner and Guthrie also contributed. During the last few years, however, other learning theorists have been making vehemently unfavourable remarks about psychoanalysis and advocating techniques of symptomatic therapy that flatly contravene most of the rules that have been widely accepted ever since Freud laid them down. The way in which Dollard, Miller and Mowrer, on the one hand, and Eysenck, Taylor and Wolpe, on the other, can derive such diametrically opposite theories of abnormal behaviour from the same Hullian type of learning theory is highly instructive. It suggests that rigorous deduction must be distinguished from plausible inference and that there is no substitute for the patient accumulation of empirical data under controlled conditions. It suggests also that some 'halo' and 'horn' effects have been at work. Freud said so many things that it is highly improbable *a priori* that all of them are true or that all of them are false. What seems to have happened is that learning theorists have picked on certain elements of what psychoanalysts say, found these elements acceptable or unacceptable, and consequently concluded that the bulk of psychoanalytic theory must be on the right or, as the case may be, the wrong track.

The title of this book of Eysenck's must not be taken too seriously. This is a pity, since many criticisms will probably stem not from the intrinsic value of the contents but from the way in which the contents disappoint the broad expectations that the title invites. The book is really volume v in a series of periodical summaries of research carried out at the Maudsley Hospital's Psychological Department. The latest phase of this research has concentrated on measurable differences between introverts and extroverts, including sufferers from anxiety states and from hysteria respectively. It has been inspired by aspects of Pavlov's progressively changing theory of personality types and by Hull's theory of inhibition. But it does not make any contact with the work of Pavlov and others on experimental neuroses, with the voluminous studies of particular human syndromes reported in Pavlov's *Conditioned Reflexes and Psychiatry* or with the substantial body of experimental work on anxiety as an acquired drive, on conflicts and on reactions to frustration that has grown out of Hull's ideas. It does not tell us much about the reasons why anxiety neurotics and hysterics differ from other introverts and extroverts, and it offers only a few perfunctory, *ad hoc* and admittedly speculative suggestions respecting the reasons why some individuals have one set of symptoms and others have a different set.

Not only is there no attempt to cover all aspects of the field denoted by the title, but the policy of using the book to report all the Department's major research projects that have not been written up in previous books produces a not-too-satisfactory compromise between a unified and systematic coverage of one topic

and a series of relatively independent papers on different topics. There is a tendency to dwell in excessive detail on matters that are not directly relevant to the main theme of the book. They may be directly relevant to matters that are directly relevant to the main theme. But that is not the same thing.

The crux of the book lies in the hypothesis that excitation is generated more rapidly and inhibition more slowly in introverts than in extroverts. It was first put forward by Eysenck some years ago with the support of Franks's experimental findings, which showed that eye-blink conditioning occurred more rapidly and extinguished more slowly in introverts than in extroverts. The reaction in many quarters was one of scepticism. For one thing, it was felt that the experiments should be repeated with other conditioned responses, in order to verify that the findings are not confined to the blinking response and that there is such a trait as general conditionability. Secondly, there are many factors that could produce differences between individuals in speed of conditioning and of extinction. They include, besides the rate of growth and asymptote of inhibitory potential (which Eysenck favours), drive level, amount of reinforcement and stimulus–intensity dynamism. Variations in any of these might conceivably account for the observed differences between introverts and extroverts. Moreover, a demonstration that introverts and extroverts differ in one of these factors would not exclude the possibility of their differing in one or more of the others also. The main distinguishing characteristics of inhibitory potential are that it accumulates with prolonged evocation of a response and that it dissipates in the course of a rest period. Many readers of Eysenck's original papers on the topic wished, therefore, that data comparing reminiscence after, say, a 24-hour interval had been secured from Franks's subjects.

The hypothesis that a higher level of anxiety drive is what makes introverts acquire eye-blink conditioning more readily is in line with what Spence and his associates have been writing. It is a tempting hypothesis in view of the concentration of both Franks and Spence on noxious rather than rewarding unconditioned stimuli and of the fact that neurotic introverts are characterized by 'manifest anxiety' symptoms. Eysenck criticized Spence's position, and he reiterates this criticism in the present book, on the grounds that some neurotics, i.e. hysterics, show slower-than-average conditioning. But this would contradict Spence's hypothesis only if it could be shown that neurotics, including hysterics, have a higher drive level than non-neurotics, a speculation that Eysenck has tentatively offered from time to time but has never actually demonstrated.

The two pieces of research that would answer the most persistent sceptics have still not been carried out. But in the meantime, Eysenck has marshalled a most imposing collection of findings, from such varied fields as perception, decrements in prolonged motor performance, mental fatigue and drugs, all of which point in the direction of a greater proneness to inhibitory processes in extroverts than in introverts. He freely admits that each of these studies by itself may be inconclusive and points out flaws in many of them. But the net effect of all these data is to win a serious hearing for his hypothesis that inhibition holds the key to the introversion-extroversion continuum.

Throughout the book, Eysenck submits his conclusions with considerable modesty and humility. He is generous to those whose work he is attacking, except for the psychoanalysts. He reiterates that he does not regard his theory of individual differences as established. He merely claims that it will yield enough testable predictions to be scientifically respectable and declares his willingness to abide by the verdict of future research. It must be admitted that he anticipates and acknowledges most of the criticisms to which the material in his book may be open.

In short, this is an important book that all experimental psychologists, especially if they are interested in learning theory, all personality theorists, all clinical psychologists and all psychiatrists ought to read. It reports many experimental results and theoretical innovations that are of interest in themselves, quite apart from their bearing on the author's theory of individual differences. The ideas it introduces certainly deserve to be thrown into the arena to hobnob with those that have hogged the limelight hitherto.

D. E. BERLYNE

An Introduction to Psychopathology. By D. Russell Davis. London: Oxford University Press. 1957. Pp. vii+388. 30*s*.

This is an attempt to bring together the disciplines of experimental psychology and clinical psychiatry. The author's training fits him well for this task: he has spent much of his time in the Psychological Laboratory at Cambridge while holding a Consultant appointment at the Fulbourn Hospital. A synthesis of this sort is always an ambitious undertaking: Masserman, Hunt, Dollard and Miller, Mowrer and others have all made similar ventures with varying success. It cannot be said that Dr Davis has been any more

successful than those just mentioned, for he has imposed on himself certain strategic limitations that give rise to an unduly rigid and doctrinaire treatment of the subject. In the first place he has entitled his book 'Psychopathology', thereby committing himself to a one-sided approach to mental illness and mental disorder which forces him to minimize the role of physical and genetic factors to an absurd degree. Furthermore, in giving a rather special meaning to the term psychopathology, he adopts an unduly mentalistic attitude both as regards the phenomenology and the aetiology of the various states described, besides perpetuating a physical-mental dichotomy that is surely by now out-dated. But Dr Davis adds to his difficulties by other defects. He has made little effort to organize his material systematically for the benefit of his readers. At times, he appears to address the psychiatrist, uninformed in the aims and methods of the experimental psychologist; at others, he addresses the experimentalist in like fashion and, at others again, he writes for the student or intelligent layman. There is nothing in common with these various essays except that they are written by one man. They exemplify the more unfortunate consequences of too loosely committed eclecticism. Then there are a number of irrelevancies appended for no very good reason beyond the fact that they constitute a particular interest for the writer himself.

Dr Davis's main thesis is that 'psychopathology is dependent on psychology and draws from psychology the principles by which to explain the diverse and unsystematic observations made during the investigation, treatment and follow-up of patients. These principles, established and proved in the laboratory, are applied in the clinic' (p. 2). He sets out to develop this theme in two parts. In Part I, he deals with various aetiological factors in mental illness and mental defect, stressing the importance of family structure, early environment and cultural influences. He gives an account of mental development and its various disorders. Thus this part is largely a discussion of clinical entities, loosely arrayed, with special reference to the influence of 'environment' conceived along fairly conventional lines. In Part II, the author describes past and current experimental methods having to do with behaviour disorders. Invoking neo-behaviourist researches and theories, he sets out to explain mental disorders in terms of these theories. In the process, both the impeccability of the experimental procedures and the richness of the clinical material suffer a sea-change so that neither universe of observation is improved. As an experimentalist the author falls into his own trap of woolliness and as a clinician retreats into his ivory tower of objectivism. Here and there throughout the work the author alludes to the contributions of psychoanalysis. He concedes that most psychopathologists owe more to Freud than they are apt to admit. He himself does not owe much; nor does he appear to have grasped the essentials of Freud's discoveries, notably with regard to the unconscious aspect of psychological processes. It is indeed surprising to encounter (p. 32) in 1957 an objection to psychoanalysis that was abundantly refuted by Freud in 1914: that Freud's original observations were made on upper-class Viennese with sexual conflicts more overt than in present-day England! As is now so commonly found in works of this kind, the section devoted to therapy is short and quite unhelpful, confined as it is to intellectual acceptance on the part of the patient and moral exhortation by the therapist armed with a novel and intricate terminology that should impress.

It must be recognized that the difficulties inherent in a synthesizing approach to psychology and psychiatry are immense and that it will be many years before we can attain to a truly synoptic account of these two disciplines. The author must be admired for his heroic attempt to bring together in one volume such a disparate collection of facts and theories, but reluctantly it must be said that the disparity remains and confusion is increased.

W. M. MILLAR

Group Psychotherapy: the Psycho-analytical Approach. By S. H. FOULKES and E. J. ANTHONY. Harmondsworth: Penguin. 1957. Pp. 263. 3*s.* 6*d.*

Though group psychotherapy is coming to be regarded as the most promising development during recent years in the treatment of psychiatric disorders, and is being increasingly adopted in hospitals and clinics, descriptions of its theory and practice have, in this country, been mainly confined to individual papers scattered through the journals, and no clear exposition for the lay reader has existed. The authors have undertaken the difficult tasks of attempting to bring together an outline of current theory and technique in the field for the information of specialists, while at the same time making their exposition simple and clear enough to enable the educated layman to follow most of it without much previous acquaintance with psychiatric theory and jargon. The compromise has been achieved skilfully, and, though it makes for a certain unevenness, it is probable that the book has gained rather than lost in value through this need for clarity in expression. Both authors are peculiarly suited to the task they have set themselves since Dr Foulkes was the first to develop analytically orientated group therapy in this country, while

Dr Anthony has extended its original use with adults to disturbed children and adolescents, and for purposes of research into the nature and origins of mental disorder. Particularly valuable is a lengthy 'introductory survey' which traces the historical development of group therapy and the debts it owes to psychoanalysis and to various earlier studies of social interaction, such as those of Kurt Lewin, Trigant, Burrow, and Lawrence Frank. It also contrasts the theory and technique of Slavson, Wolf, the Tavistock Group and other workers with that evolved by the authors and termed 'group-analytic psychotherapy'. The discussion is admirably fair, detached and balanced, a welcome change in psychiatric fields, where workers are often prone to argument based on personal rivalries and animosities. It emerges clearly that essentially similar theories and methods have evolved from the various 'schools' of group therapy, the main difference lying in the degree to which treatment of individuals *in* a group by the doctor has evolved into treatment *of* the group *by* the group, the doctor becoming increasingly passive and interfering only when the process is blocked, or when his special knowledge is required to help them translate and understand the answers they have produced unwittingly. The authors claim to have progressed furthest in this direction, and operate on the assumption that '...the group tends to speak and react to a common theme as if it were a living entity, expressing itself in different ways through various mouths. All contributions are variations on this single theme, even though the group is not consciously aware that this is so and do not know what they are really talking about.'

The remainder of the book is devoted to the changes in therapy that this revolutionary discovery effects, and to detailed descriptions of technique in its practical application. Most of this is clearly written and full of valuable information even for those with much experience of group psychotherapy, though some of the theoretical speculation, particularly the attempt to describe the process in Gestalt 'figure-ground' terms, introduces unnecessary complication without really explaining anything. The weakest sections are the clinical illustrations. The verbatim records of children's groups are not accompanied by adequate explanations of the latent meaning of what was said, or the purpose it served, while the long description by a member of an adult group of her experience in it is unconvincing to the inexperienced reader, and would have been better replaced, or at least supplemented, by extracts from tape recordings which are readily available and can show the various phenomena encountered in groups with convincing clarity.

R. SKYNNER

Family and Social Network. By ELIZABETH BOTT. London: Tavistock Publications. 1957. Pp. xi+252. 30*s*.

This is an account of an investigation conducted by a social anthropologist, a social psychologist and two psychoanalysts into the internal and external relationships of twenty London families, with particular reference to the social roles of husband and wife. The families were chosen because they were 'ordinary' in the sense that the research workers went to find them; they did not come to the research team for help, although two of the families had approached another agency for advice. A large element of self-selection necessarily entered into the choice since the inquiry involved the voluntary collaboration of the subjects over a prolonged period. Information was obtained in the course of a dozen or so home interviews lasting an hour or more. This was supplemented, in fifteen of the twenty cases, by clinical interviews with a psychoanalyst at the Tavistock Institute of Human Relations. The present volume is concerned primarily with the social data obtained. It deals with five main topics: (i) The degree of conjugal role segregation shown by the husband and wife, i.e. the extent to which domestic and leisure activities were shared. (ii) The extent to which relationships outside the conjugal family were with a closely knit network of friends, neighbours, and workmates, or were with scattered and unconnected acquaintances. (iii) The degree of intimacy of contact with members of the extended kinship group. (iv) The views of the informants on what constitutes 'normal' behaviour in various aspects of family life. And (v) The views of the informants on questions of social class and status mobility.

The majority of informants had a relatively loose-knit network of friends and acquaintances and little contact with kin. They had difficulty in making articulate their views on social class and the norms of family life. Bott suggests that whereas, in a close-knit network, external sanctions determine attitudes and behaviour directly, in a loose-knit network people do not acquire their values solely by internalizing them from outside, but tend to project their own standards on to others. The principal conclusion emerging from the inquiry is that the more closely connected the family's social network the greater the tendency for husband and wife to have a strict division of labour in the home and to lead more or less separate lives as far as their leisure activities are concerned.

A later book by Dr A. T. M. Wilson will deal with the 'unconscious aspects of the relationship between husband and wife, of the interplay of their personalities, and of the emotional tasks of marriage'. Bott admits that the small size of the sample makes firm conclusions impossible and her findings are put forward as hypotheses which must be tested in the light of further comparative evidence. It is to be hoped that Dr Wilson's forthcoming book will justify the use of intensive interviewing techniques, because none of the topics dealt with in the present volume was outside the scope of an ordinary random-sample survey. In some respects the investigators would have done better if they had relied upon shorter interviews with a larger number of people. What they would have lost in depth would have been amply repaid by having a representative sample.

ANTHONY H. RICHMOND

The Nature of Conflict: Studies on the Sociological Aspects of International Tensions. By The International Sociological Association, in collaboration with JESSIE BERNARD, T. H. PEAR, RAYMOND ARON and ROBERT C. ANGELL. Paris: UNESCO, 1957. Pp. 314. 21*s.*

The statement is made in the Preface to this book that it presents an up-to-date survey and evaluation of research by sociologists and social psychologists into the nature, conditions and implications of human conflict, and particularly conflict between nations. The Introduction is a review of the UNESCO Tensions Project, worked out by the Central Conference, Mexico City, 1947, and of progress made since 1948.

From the point of view of psychology the interest of the second chapter, by Emeritus Professor T. H. Pear, is considerable. He has made a survey and evaluation of publications and researches which seemed to him of special importance, dealing with psychological studies of the tensions leading to conflicts. It is a fair and interesting survey and will be useful to social psychologists.

Prof. Jessie Bernard contributes a chapter on the sociological study of conflict, Prof. Raymond Aron contributes one upon conflict and war from the point of view of historical sociology, and Prof. Robert C. Angell a chapter on the discovery of paths to peace. The last writer takes the view that states which are in interaction will enjoy peace, not so much by the mere absence of war, but only if they are parts of a social system which embraces them.

The book concludes with a useful list of 1160 titles of relevant books and other publications, prepared by the 'Centre d'Études Sociologiques', and a list of periodicals. If the reader feels disappointed because he does not find in the book decisive answers to his biggest questions, it must be realized that such answers cannot yet be given, but he will find a study of many problems and ideas which will help him.

R. W. PICKFORD

Perceptanalysis. By Z. A. PIOTROWSKI. New York and London: The Macmillan Company. 1957. Pp. xvii+505. 47*s.*

Despite the wider implications of its title, Piotrowski's *Perceptanalysis* is concerned entirely with the Rorschach method. As such it invites comparison with Klopfer's *Developments in the Rorschach Technique* (reviewed in this *Journal*, 1955, **46**, 72–3 and 1957, **48**, 150–1).

Lacking experience of testing Piotrowski's rationale, one has to rely on impression which suggests that, while on occasion he shows quite brilliant clinical insight, the work is far below that of the Klopfer group as regards scientific approach. There are numerous cases of *non sequitur*, circular reasoning, and the like, and—as is unfortunately all too common in Rorschach literature—statements are often made quite dogmatically without supporting evidence. One of the few places where detailed figures are quoted is in a discussion of the ranking of plates in terms of initial reaction time; here the issue is confused by the figures being mean ranks whereas the discussion refers to mean reaction times. The interpretation seems to be at variance with the figures, but there is no way of telling whether this is indeed the case.

It is difficult to get an all-over picture of Piotrowski's scoring system, since he gives no summary list of symbols, ratios calculated, etc. Two important innovations, however, stand out. First, his admission of an *MC* response, i.e. one that contributes equally to the human movement and the colour scores, and an extension of this principle to other combinations of two or even three determinant categories. This allows him to abolish the necessity to recognize 'precedence' of one determinant over another, and the often meaningless distinction between 'main' and 'additional' scores of the Klopfer method. Secondly, his treatment of shading responses is entirely original, in that he uses two basic symbols only: *c* for

'light grey and variegated shading' and *c′* for 'dark-shading'. Each of these is also made the subject of a weighted score analogous to 'colour sum'. While this treatment obviates many scoring problems, it fails to recognize the interpretative significance of the texture response which many Rorschach workers find particularly illuminating; the author also weakens his case by the remark (p. 257) that 'Percept-analysis is not so precise a measuring instrument for a simplification of this degree to cause significant errors'. In this connexion, too, Piotrowski produces the following 'gift' to those who hold Rorschach 'jargon' up to ridicule: 'As the pig is a very active animal it also symbolizes the *c′R* tendency while the sheep, which is much less active, symbolizes the *cR*. Liddell reported that Pavlov did not think that one could develop an experimental neurosis in pigs...(but)...succeeded where Pavlov had failed and found that pigs become increasingly aggressive and violently active as they are made more and more uncomfortable by the conditioned reflex experiments. Thus, they behaved in a typical *c′* manner. Sheep on the other hand became less and less active, more stiff, and developed cardiovascular symptoms as the experimental neurosis grew stronger. Thus, their behaviour was of a *c* type. In order to alleviate anxiety, the pig increased its activity, while the sheep decreased its activity. A human being, having a more complex central nervous system than either pig or sheep, is capable of both types of reaction.'

A comparable innovation, on the interpretative level, is the recognition of specific 'shock' elements associated with each plate. In the reviewer's experience, plate IX frequently produces disturbances in response which appear to be due to difficulties in form interpretation rather than to the colour content. To describe such a disturbance simply as 'Plate IX shock' (for which, of course, the author has a rationale) would appear to be an advance on the assumption—which one sometimes suspects has been made—that because stimulus variables are present they are necessarily operative.

Notwithstanding some of the foregoing criticisms, a fair summing-up would be that this is a stimulating book, with many original and controversial features to assure its welcome. Particularly interesting is the information regarding the background and early history of the technique, including frequent references back to Rorschach's own writings. Only the out-and-out sceptic will deny that it throws fresh light on the many problems of personality dynamics.

B. SEMEONOFF

Insight: A Study of Human Understanding. By B. J. F. Lonergan. London: Longmans, Green. 1957. Pp. xxx+785. 63*s*.

A long and closely argued philosophical book by the Professor of Dogmatic Theology in a Catholic University is not meat for many psychologists. And in Father Lonergan's book there are lengthy meta-physical stretches which probably no psychologist *qua* psychologist could be troubled to read. Never-theless I found, and I think others will find, much in it to reward them. My admiration increased as I read, and my final impression was that this is a great book, and a book of which psychologists ought to know.

It is about Insight, the activity of the organizing intelligence. It examines the nature of Insight, taking examples in the first instance from mathematics and physics, with which subjects the author shows himself to have a considerable mastery; it analyses the insights of common sense in a most useful couple of chapters; it complements the concept of Insight with the concept of Oversight, and discusses the biases of the understanding and the implications of the flight from understanding. It goes on to discuss the philosophical implications of understanding. The author himself summarizes his theme as follows: 'Thoroughly to understand what it is to understand, and not only will you understand the broad lines of all that is to be understood, but you will possess a fixed base, an invariant pattern, opening upon all further developments of understanding.'

A book with such a theme may well be regarded as just another instance of the delusion of the omni-potence of thought, and dismissed forthwith by the scientifically minded psychologist. In this case I think he would have made a mistake. This book shows a sustained level of first-hand philosophical thinking of a high order. It combines contemporary scientific knowledge and an awareness of contemporary problems with a broad grasp of philosophical tradition. It is cogently reasoned and lucidly written. It is often penetrating and humane in its observations. It can be recommended to those psychologists interested in cognitive processes who have time and patience, a tolerance for philosophy, and a spare three guineas—for this is a book which one should possess for oneself if one is to trouble with it at all.

L. S. HEARNSHAW

The Collective Dream in Art. By WALTER ABELL. Harvard: University Press. London: Oxford University Press. 1957. Pp. xv + 378 + 39 Plates. 60*s*.

This is a valuable book deserving serious attention from anyone interested in the function of art and of artists in society and willing to reflect on the possibility of integrating historical, sociological and psychological concepts with the findings of art history in order to understand that function. The author's view, succinctly stated on the inside book cover, is that: 'Arts, myths and other forms of cultural expression are to the society that produces them what dreams are to an individual, i.e. symbolical projections of underlying mental states caused, in turn, by the outward circumstances of life.' Walter Abell, who died before his book appeared, was Professor of Art at Michigan State University. He wrote as a specialist in the history of the visual arts, seeking both to enrich his own field by borrowing from allied disciplines and also to suggest an integrated viewpoint about which he clearly felt deeply.

The author's own discipline has contributed to two aspects of the book distinguishing it pleasantly from the crude psychoanalytic ruminations that psychologists are sometimes asked to accept from artistic and literary colleagues. First, Prof. Abell relates his theory to a specific problem in the history of art, namely that of the grotesque in Gothic sculpture. To this he devotes more than a third of the volume in the spirit of trying out a hypothesis to see if it fits the facts. Secondly, Prof. Abell has been chary of accepting either psychological or historical explanations that claim to be all-inclusive. Though, for example, his psychological borrowings are mainly from Freud and Jung, he steers clear of the biological and cultural narrowness that colours some extreme forms of psychoanalytic orthodoxy. Similarly, when writing of the 'collective dream', Professor Abell does not seem to have lost himself in the metaphysical railway cloakroom of the Jungian collective unconscious nor taken over the curious archetypal luggage supposedly deposited there. Again, while borrowing from history and sociology the view that we should seek in economic and social conditions a partial explanation of the negative, of the positive and of the neutral or accepting attitudes he believes to be externalized in art, Prof. Abell is well aware of the fallacies in Marxism. His approach therefore is critical and in these respects balanced. Whether or not Prof. Abell establishes a case remains a matter for discussion. For example, is his viewpoint really new? Has he, as he himself wonders, relied too heavily on direct projection and repression and allowed compensatory mechanisms too little weight? Is his attempt to balance socio-economic and personal factors in the production of art forms adequate or subtle enough? How many artists would accept his discussion of their role, sympathetic as it is? Though some incidental comments are extremely good, the exposition, especially in Part I, strikes the psychological reader as somewhat laboured. Throughout, though less in later chapters, this reviewer sighed periodically for the sparkle with which a Huizinga or a Paul Hazard would have handled comparable material. Occasionally Prof. Abell seems to write with elation about his approach and provokes the reader to be as chary as the author can be of the easily all-inclusive.

But when all this is said, one remains with the impression of a writer, wide in interest, struggling to be objective, unwilling to tolerate cheap negation, and in love with the arts and honesty even if experience is tragic. This is an honourable impression to have left behind and one can but write of it with respect. The book is well produced and includes more than a hundred photographs. These succeed in providing appropriate illustrations, but the presentation would be more effective if in a second edition the publishers could increase the size of the individual photos and arrange them where relevant in the texts. As it is the book is an asset to any good library.

JOAN WYNN REEVES

Instinct in Man: In the Light of Recent Work in Comparative Psychology. By RONALD FLETCHER. London: George Allen and Unwin. 1957. Pp. 348. 40*s*.

This is an exposition and critique of modern views on instinct, with particular reference to the contributions of ethology and psychoanalysis. Dr Fletcher has evidently read widely in both fields and has thought closely about the theoretical issues involved. Unfortunately, he appears to be without personal experience in either discipline and his exposition in consequence lacks that sense of reality which direct exposure to the facts alone can bring. None the less, his work will be of considerable help to students of psychology and others seeking guidance in this confused and controversial field.

O. L. ZANGWILL

Statistics for the Social Sciences. By T. G. Connolly and W. Sluckin. London: Cleaver-Hume. Second ed., 1957. Pp. viii + 166. 16*s*.

Everyone who has been concerned in the teaching of statistical method to the mathematically unsophisticated has probably, at some time or other, thought of writing his own text. Only in this way can one be sure that the student has before him, in black and white, everything that he ought to have—and as little as possible to frighten him. Failing this, such a teacher could do worse than fall back on Drs Connolly and Sluckin, who have had a great deal of relevant experience. Their exposition is usually clear, and all the basic techniques are adequately covered.

It is easy to criticize a book of this sort on the basis of what, beyond the essentials, one would oneself have included or omitted. In the present instance, one might question the necessity for more than a passing mention of the mean deviation; the retention of the Spearman 'footrule' in addition to *rho*; and the desirability of burdening the beginner with the labour of computing an approximate tetrachoric correlation coefficient by solving an equation involving the first two powers of r_t. On the other hand, one would perhaps have liked to see included a treatment of point biserial as well as biserial correlation, of the contingency coefficient, and possibly even of two-way problems, as well as one-way, in the new chapter on the analysis of variance. However, one cannot have everything.

It is a pity that a few misprints have been overlooked, including at least one in a formula which is quoted once only. Again, a distinction might have been made between tables which are merely incidental to worked examples, and those to which reference must be made, e.g. in applying tests of significance. The latter are well chosen, and present the required information concisely and economically. Finally, one could wish to have certain examples (e.g. correlation from a scattergram) worked out more fully, while others of more purely theoretical interest, such as the derivation of values of ordinates of the normal curve, would perhaps have been better omitted.

B. Semeonoff

Ergonomics: Human Factors in Work, Machine Control and Equipment Design. Vol. 1, No. 1. Nov. 1957. London: Taylor and Francis. Pp. 100. 25*s*.

The sampling error that might result from attempting to review a journal on the basis of its first number makes it unwise to do more than welcome a new contemporary and wish it success. But one of the attractive features of this, as a first number, is that it comes into the world with so little fuss. There is a short editorial defining the field, followed by eight workmanlike papers which are so much at home with one another as to give the impression that we are reading a journal whose scope and policy have long been established. In this number the authors are well distributed between the U.S., the U.K., and the Continent, and the papers range from concrete practical studies such as 'Human Limitations and Vehicle Design' by Ross A. McFarland, or 'Factors in Fatigue and Stress in the Operation of High-Speed Diesel Passenger Railway Cars with only One Driver Present' by Robert S. Schwab, to an ingenious mathematical approach to the estimation of skill by J. R. De Jong. Between these extremes there are studies of working conditions, of training methods, and techniques of validation in terms of user preference. All in all, this is an excellent beginning, and there seems every reason to predict a long and successful future.

James Drever

OTHER PUBLICATIONS RECEIVED

More About the Backward Child. By HERTA LOEWY. London: Staples. 1957. Pp. 138. 15*s.*

Fear: Contagion and Conquest. By J. C. MOLONEY. New York: Philosophical Library. Pp. 140. $3.75.

The Substitute Teacher as Indicator. By M. KOSKENNIEMI. Helsinki: Finnish Academy of Science and Letters. 1957. Pp. 58. 150 F.mk.

A Course in Multivariate Analysis. By M. G. KENDALL. London: Charles Griffin. 1957. Pp. 185. 22s.

Self-conditioning: A New Way to a Successful Life. By HORNELL HART. London: Oldbourne Press. 1957. Pp. 198. 15*s.*

The Child Within the Group: An Experiment in Self-government. By M. E. TURNER. Stanford University Press. London: Oxford University Press. 1957. Pp. 93. 24*s.*

Stammering and its Treatment. By FRANKLIN BROOK. London: Pitman Medical Publishing Co. 1957. Pp. 134. 21*s.*

Stuttering: A Psychodynamic Approach. By DOMINICK A. BARBARA. London: Hutchinson. 1958. Pp. 255. 25*s.*

Fun Comes First for Blind Slow-learners. By MILDRED B. HUFFMAN. Illinois: Thomas. Oxford: Blackwell. 1957. Pp. 157. 37*s.* 6*d.*

Rencontre: Contributions à une Psychologie Humaine Dédicés au Professeur F. J. J. Buytendijk. By VARIOUS AUTHORS. Utrecht: Spectrum. 1957. Pp. 520. 25 H.fl.

Pupils' Arithmetical Ability: Differences due to the Teacher. By T. VAHERVUO. Helsinki: Finnish Academy of Science and Letters. 1957. Pp. 48. 200 F.mk.

Introductory Psychology: An Approach for Social Workers. By D. R. PRICE-WILLIAMS. London: Routledge and Kegan Paul. 1958. Pp. 203. 18*s.*

Perceptual Development: An Investigation within the Framework of Sensory-Tonic Field Theory. By S. WAPNER and H. WERNER. Clark University Press. 1957. Pp. 95. No price given.

A Study of the Major Psychoses in an English County. By M. SHEPHERD. London: Chapman and Hall. 1957. Pp. 159. 25*s.*

Leader Attitudes and Group Effectiveness. By FRED E. FIEDLER. University of Illinois Press. 1958. Pp. 69. $1.75.

Value in Social Theory: A Selection of Essays on Methodology. By GUNNAR MYRDAL. Ed. by P. STREETEN. London: Routledge and Kegan Paul. 1958. Pp. xlvi+269. 32*s.*

Emotional Conflict. By PETER FLETCHER. London: Pan Books. 1958. Pp. 158. 2*s.* 6*d.*

Prescription for Survival. By BROCK CHISHOLM. Columbia University Press. London: Oxford University Press. 1957. Pp. 92. 20*s.*

Mental Deficiency. By J. E. W. WALLIN. Brandon: Journal of Clinical Psychology. 1956. Pp. 200. $5.00.

The Uses of Argument. By STEPHEN TOULMIN. London: Cambridge University Press. 1958. Pp. 264. 22*s.* 6*d.*

Bureau Ednl. Psy. Research
...ING COLLEGE

ERNEST JONES, M.D., D.Sc., F.R.C.P., 1879–1958

OBITUARY NOTICE

ERNEST JONES, M.D., D.Sc., F.R.C.P., 1879–1958

On 11 February 1958 Ernest Jones died in London at the age of 79 years. With his death was severed the last link with that classical period in the development of psychoanalysis when, with the assistance of a small but devoted band of acolytes drawn from Vienna, Zurich, Budapest, Berlin, London and New York, the discoveries of Freud were spread abroad, and psychoanalysis, from being a science with for ten years but one solitary advocate, became a firmly organized movement centred on an International Association. Ernest Jones was the last survivor of this band of conquistadors. In his time he led an ever-multiplying contingent of English-speaking analysts through Canada, New York and London to the four quarters of the British Commonwealth; fought in every battle of offence and defence throughout the pioneering period that lasted almost twenty years (1908 to 1928); and on the death of Freud in 1939 became the acknowledged leader of psychoanalysis throughout the world. During that time he organized and expanded Freud's approaches to the cognate sciences of psychology, sociology, and anthropology, and to the territories of aesthetics, movements which were regarded as unwarranted invasions. And he left behind him a richly documented record of progress as a teacher, research scholar, writer, organizer and finally leader in the field of psychoanalysis.

Born in 1879 in Gowerton, Glamorgan, of Welsh parentage, he proceeded from the village school, where he was recognized as a boy 'of parts', to Swansea Grammar School, thence to University College, Cardiff; and, having later studied at University College Hospital, obtained the degrees of M.B., B.S. at the University of London in 1901. In 1903 he took the senior degree of M.D. (Lond.); in 1904 he became an M.R.C.P. and in 1905 obtained the D.P.H. (Cambridge). With both junior and senior medical degrees he was awarded gold medals. He was elected Fellow of the College in 1942, and in 1954 was awarded an honorary D.Sc. by the University of Wales.

Jones's hospital career had a significant bearing on his ultimate choice of profession. Like many brilliant students who cannot see their way clear to follow the conventional track to Harley Street and the practice of consulting medicine, he lingered in hospital work, taking the usual house appointments in U.C.H. He was also clinical assistant at the Royal Hospital for Sick Children and the Ophthalmic Hospital. More significant were his appointments as clinical assistant at the National Hospital, Queen Square, and as registrar and pathologist at the West End Hospital for Nervous Diseases. These were followed, after a few intermediate hospital posts as assistant physician and physician, by a lectureship in practical neurology at the London School of Clinical Medicine. So that already before he left England for Canada, his interest was fully slanted to neurology, to which speciality he had already contributed some papers (published in *Brain*). In this respect there was a certain parallel between his progress and the early orientations of Freud.

It was, indeed, because of this interest that he went to Munich for post-graduate study in neurology and psychiatry and became acquainted with the writings of Freud. His

absorption in the new science increased when in 1908 he went to Toronto as demonstrator of medicine and psychiatry at the University and psychiatrist to the Toronto General Hospital. Later he was appointed Professor of Psychiatry at Toronto and Director of the Ontario Clinic for Nervous Disorders. By this time he was already employing a psychoanalytic technique in his work and published some of his results in the *Journal of Abnormal Psychology* in 1909.

During his stay in Canada psychoanalysis took up an increasing amount of his attention. With his active support it had already obtained a footing in New York. And when the International Psychoanalytical Association was founded in 1910 he was one of its foundation members. The die was soon to be cast. Encouraged by Freud whom he met during the latter's visit to America in the company of Ferenczi he decided to return to England, and in 1913 did so.

At first a lone wolf, as Freud in the late Victorian nineties had been a lone wolf, Jones was to find support through the accident of international circumstances. The psychological casualties of the First World War had provoked the interest both of psychiatric specialists and of some academic psychologists. The more discriminating amongst these were also stimulated by Freud's theories of the origin of the psychoneuroses. And so after reconnoitring the situation in London for some time Jones found opportunity and support to set up a London Psychoanalytical Society. The future of psychoanalysis in England seemed assured.

But, alas for the prospect, psychoanalysis, being based on the psychology of unconscious conflict, is itself continually prone to fission between its adherents. It soon became apparent to Jones that the members of the newly formed Society were not at one on matters of principle; and on matters of principle Jones was always adamant. So in 1919 he formed the British Psychoanalytical Society which met at first in his consulting room and later in that of Dr Riggall in Wimpole Street. A small group, most of whose members have since died, met together and by a process of co-option and apprenticeship slowly built up the British Branch of the International Psychoanalytical Association, of which latter, incidentally, he was about to become President in 1920, the year in which he founded the *International Journal of Psycho-Analysis.*

Thenceforward the progress of the British Society was increasingly rapid and by 1926 the articles of association of the Institute of Psychoanalysis were signed. The Institute gave formal cover to the British Society, took over the *International Journal* which Jones edited from its inception until 1939, published in conjunction with the Hogarth Press the *International Psycho-Analytical Library*, which under Jones's editorship produced some fifty volumes of mostly standard works on psychoanalysis, undertook under the auspices of its Training Committee (of which Jones was Chairman up to the Second World War) the teaching and training of psychoanalytical students, organized the 'orientation courses' for the medical profession and the general public, and founded the London Clinic of Psychoanalysis of which Ernest Jones was the first Director.

In spite of these many activities, opposition to psychoanalysis in Britain, as elsewhere, remained unabated, was indeed increased owing to the activities of many unqualified persons who called themselves 'psychoanalysts'. Medical uneasiness increased to such a pitch that an investigating Committee was appointed by the British Medical Association. Jones was the sole psychoanalytical representative on a body of about twenty, most of whom were not conspicuous for their support of psychoanalysis. This was Jones's last

gladiatorial display. Not to put too fine a point on it, he fought his critics on the committee, not to mention numerous hostile witnesses, to a standstill; with the result that the B.M.A. Committee refrained from expressing any collective opinion, but laid it down that 'the term "psychoanalysis" can be legitimately applied only to the method evolved by Freud and to the theories derived from the use of this method'. As far as Britain was concerned this was a psychoanalytical charter and definitely marked the close of the pioneering phase.

From this time Jones, who had been President of the International Psychoanalytical Association from 1920 to 1924, played an increasing part in the international policies of psychoanalysis. After the death of Abraham of Berlin he fell heir to the position of leader, was re-elected President of the Association in 1932, a position he held until 1949, when, on retiring, he was elected with acclamation its honorary President. Amongst his many achievements in this field the most important was his conduct of negotiations with the American Psychoanalytical Association which helped materially to maintain the international unity of psychoanalysis.

But these biographical items, impressive as they are, do not convey adequately the status of Ernest Jones as a scientist, as a leader or as a man. As a leader the foundations of his psychoanalytical fame rest firmly on the facts, first, that he defended psychoanalysis during that vulnerable period when it was subject to continuous, violent and often vituperative opposition; second, that through his unremitting efforts he was able to preserve psychoanalysis from internal disruptive influences. Jones had unique qualifications for these tasks. Keen and nimble of wit, he could use his well-stocked mind with devastating effect on those opponents who crossed swords with him. He was a born controversialist with a sharp tongue and a quick eye for the weaknesses of an argument; and he had no mercy on opposition that was based on malice. His counter attacks were not, however, purely polemical in character. He was an accurate, polished and on occasion persuasive exponent of psychoanalysis; having cleared aside the debris of controversy, he could improve the occasion by giving lucid presentations of even the most recondite psychoanalytical theories. The same positive qualities were shown in his leadership. As a Congress Chairman he could with a few trenchant phrases reduce to order the bear-garden confusion of a continental discussion, and at the same time persuade opposing parties to however reluctant compromise.

As a scientist Jones's position was unassailable. Thoroughly versed in psychoanalytic method and at the same time closely conversant with the methodologies of general psychology and sociology as well as those of natural science, immensely industrious, with a quick perception of essentials and a scholar's grasp of their implications, he made an ideal research fellow, and having above all an orderly, even meticulous mind, he was able to dovetail his contributions tidily with the existing structure of both psychoanalytical and psychological theory. Although a great deal of his earlier work was inevitably didactic, he broke new psychoanalytical ground in many directions both theoretical and clinical. The various editions of his *Papers on Psycho-analysis* combined the virtues of a good text-book with the stimulating qualities of a research manual. His work was always accurate and well documented, and although not so profound as that of his master Freud amply compensated for this by its comprehensiveness of range.

But valuable as these contributions were it was in the field of applied psychoanalysis that he gained a reputation amongst the cultivated and discriminating public. Here again, his range of interest was extremely wide and gave him ample opportunity to display his almost obsessive erudition. Always an encyclopaedic reader, with an insatiable curiosity and a quick appreciation of relevant detail, his psychoanalytical contributions to literary and artistic problems and puzzles, to folklore and myth, to various aspects and manifestations of religion, to linguistics and a host of other social, cultural and anthropological subjects were always illuminated by a wealth of apposite instance. Amongst literary critics he will best be remembered by two of his works, his essay on Hamlet and his biography of Freud. He was one of the outstanding exceptions to the general rule that psychoanalysts, although expert to the point of redundancy in building up case-histories, are of their professional nature not very good (that is to say, unbalanced) biographers. Jones's lifelong concern with the unconscious mind did not professionalize or blunt his humanistic perceptions and values. The Freud biography alone, an immense compendium of fact embroidering a sustained characterological and ideological study of one of the world's great men, would be sufficient to secure him a lasting reputation amongst the discerning.

But professional achievement alone, though justifying fame, is not the ultimate criterion of greatness. It is, indeed, a commonplace that truly distinguished men give witness of their distinction in their private as well as in their professional lives. And it can be said without reservation that Ernest Jones was a man of distinction. Although this is a quality that is hard to pin down, one cannot, therefore, refrain from speculating on the personal sources of his professional eminence.

Now the secret of the careers of men lies somewhere between their creative drives and their defences against the more unruly of their passions. Jones was a Welshman born and bred, passionate in temperament, sensitive in feeling and perception, combative as well as impatient in spirit and in his own way endlessly ambitious. Moreover, as a boy he had absorbed the traditions of success that tend to prevail in impoverished and provincial principalities and which have driven the Welsh to frequent eminence in the overlord kingdom of England. Equipped by the accident of birth with good brains he learned at an early age to discipline those more sentimental Celtic accesses that would have impeded his advancement. And it was soon clear that he would achieve success in any career he elected to pursue. All that was lacking was a calling. Neurology indeed tempted him mainly through its demands on his intelligence and on his love of complex problems; but it was not until he fell under the sway of Freud that he developed a consuming passion that could sublimate his more emotional urges.

By the same token psychoanalysis afforded him a disciplined outlet for an unusual degree of filial piety. His scientific allegiances to Freud were immensely reinforced by a personal devotion, which, although it never erred on the side of idolatry, provided him with a private ego-ideal which ultimately developed into an identification. He could not himself have been a Freud, but he could at least conduct, embellish and spread the ideas of his Homeric mentor, and he could at least render tribute to Freud's greatness by every expedient of scientific defence, including the preservation of psychoanalysis from internal erosion and external bowdlerization. It was this urge incidentally that accounted for his

disregard of unpopularity. The well-being of psychoanalysis meant more to him than success in professional diplomacy.

Not only so. It was this devotion to a Freudian ego-ideal that governed his private life and relations. There were not many who were admitted behind the curtain of his professional devotions. Those who were accorded this privilege found him simple, passionate and human, not disdaining the frailties but with few of the foibles of common life. For above all else Ernest Jones was a realist. Indeed, the ultimate secret of his devotion to psychoanalysis was that it enabled him to tie his passion for realism to psychobiological moorings. In doing so he set a standard not only for himself but also for those who follow in his professional footsteps.

EDWARD GLOVER

VISUALLY CONTROLLED LOCOMOTION AND VISUAL ORIENTATION IN ANIMALS*

By JAMES J. GIBSON

Psychology Department, Cornell University

A general theory of locomotor behaviour in relation to physical objects is presented. Since the controlling stimulation for such behaviour is mainly optical, this involves novel assumptions about object perception and about what is called 'visual kinaesthesis'. Evidence for these assumptions is cited. On the basis of this theory it is possible to suppose that animals are visually oriented to the surfaces of their environment, not merely to light as such. In short, it is possible to explain why they seem to have space perception. Implications of this approach for maze-learning are pointed out.

I. Introduction

The locomotion of animals seems to be controlled by optical as well as by other kinds of stimulation for, in many species, locomotor behaviour ceases when optical stimulation fails, as in total darkness. In bats, to be sure, we now know that locomotion occurs in darkness under the guidance of auditory stimulation, or echolocation (Woodworth & Schlosberg, 1954, 359 f.), but this interesting discovery only serves to emphasize the importance of vision for the majority of animals in the business of getting about the environment, of avoiding obstacles, and of seeking food. Students of animal behaviour have long been concerned with locomotion at the level of reflexes and forced movements, but not at the level of averting collisions or of approaching goals: there is a considerable literature on phototropisms, taxes, and kineses (e.g. Fraenkel & Gunn, 1940), but almost nothing on such acts as the pursuit of prey or the flight from a predator. The reason for this failure is probably connected with the generally accepted assumption that the ability of an animal to respond to *light* and the ability of an animal to respond to *objects* are quite different problems. The first is supposed to be physiological and the second psychological. A tropism implies only a sensory process, but a response to an object at a distance implies a perceptual process and this is much more complex. The physiologist is willing to study the former but he is glad to turn over the latter to the psychologist. The problem of how we respond to objects at a distance has been, in truth, a puzzle for psychologists ever since Bishop Berkeley. Depth or space perception has been taken to be an internal process involving the interpreting or organizing of sensations. Theories of this process have never achieved either simplicity or plausibility and, if it is a process of association, the difficulty arises of why animals do not make more errors than they do in the acquisition of locomotor behaviour. The writer has argued (1950) that the only way out of this theoretical muddle is to discard the original assumption about sensory response to light on the one hand and perceptual response to objects on the other. Perhaps the distinction was based on a mistaken conception of vision in the first place. If the theories of space perception do not provide an adequate explanation for the visual control of locomotion in animals, let us

* The first draft of this paper was written in 1955–6 at Oxford University under a Fulbright award. The author wishes to thank Prof. George Humphrey, members of the Institute of Experimental Psychology, and others at the University for kindnesses beyond the demands of academic courtesy. He also thanks these and many other investigators visited during the year for cordial and stimulating discussion.

forget about the former and pay heed to the latter. Locomotion is a biologically basic function, and if that can be accounted for then the problem of human space perception may appear in a new light. The question, then, is how an animal gets about by vision. How does it react to the solid surfaces of the environment without collision whenever there is enough light to see them by? What indicates to the animal that it is moving or not moving with reference to them? What kind of optical stimulation indicates approach to an object? And how does the animal achieve contact without collision? What governs the aiming and steering of locomotion? (The problem of 'roundabout behaviour' in the presence of a 'barrier', as it appears in the writings of Lewin (1933, 1935) and later Hull (1938, 1952) is a different question. It is as close as psychologists have come, however, to a psychology of locomotion.)

II. POSTULATES FOR A GENERAL THEORY OF THE VISUAL CONTROL OF LOCOMOTION

The starting point for a theory of locomotion in the higher animals would be the fact that they have evolved eyes (Walls, 1942) which can register not merely light but the objects of an illuminated environment. The optical basis for this assertion has been accumulating for years but the conclusion has not generally been drawn despite its having radical implications for a theory of behaviour. The function of a highly developed mosaic of visual receptors is to respond to what might be called the *projective* capacity of light. When light is many-times reflected in all directions from an array of surfaces—when it 'fills' the environment as we say—it has the unique property that reflected rays will converge to any point in the medium. The objective environment is projected to this point. If an eye is placed at this point, it can register a sector of the projection by the familiar process of the formation of an image. The central hypothesis of the theory to be presented is that the *patterns* and the *changes of pattern* of this projection are stimuli for the control of locomotion relative to the objects of the environment. For the sake of explicitness, it might be well to list the assumptions singly. The first six assumptions deal with the projection of light to the position of a motionless animal, that is, with the optic array to a stationary point. The remaining postulates assume an animal in motion and deal with the optic array to a moving point.

(1) *The locomotor environment of animals consists of the surfaces or interfaces between matter in the solid state and matter in a liquid or gaseous state. This generalization holds for aquatic, terrestrial, or aerial environments, and for animals at various phylogenetic levels.*

(2) *Solid surfaces generally reflect rather than transmit light. They also generally prevent rather than permit locomotion. Consequently, the remaining liquid or gaseous regions of the environment constitute a* medium *both for the transmission of light and for the movements of organisms.*

(3) *If the environment is illuminated there will occur a many-times reflected flux of light in the medium. If the solid surfaces and their parts vary in reflectance a projection of these differential reflectances is obtainable at any point in the medium.* That is, the rays converging to this point will have different intensities (and frequency compositions) in different directions. They constitute what may be termed an* optic array.

* This is a simplified statement, holding for an ideally diffuse illumination. In other conditions of illumination, what we loosely call 'shadows' as well as reflectances must enter into consideration. The extent to which animals can discriminate in their behaviour between shadows and solid surfaces is a complex problem which can here be omitted.

(4) *An eye is a device which registers some portion of the 'pattern' or 'texture' of an optic array to a station point. Conversely, the pattern of an array is a stimulus for an eye.*

This postulate needs explanation. Eyes, to be sure, are of different types in different animals (Walls, 1942). At one extreme, a pair of eyes can include nearly the whole pattern of a 360° optic array, each eye registering an approximate hemisphere with little overlap. This is the case with fish. At the other extreme, a pair of eyes can register overlapping sectors or cones of the frontal optic array, each picking up the same pattern from a slightly different station point. This is the case with primates in whose case the eyes are mobile and co-ordinated, being capable of exploratory (saccadic) fixations and convergences. The total array can then be registered only by rotating the eyes and head. The registration process is successive, not simultaneous, since different angular sectors of the array are picked up at successive moments of time. Nevertheless, by a mechanism as yet not well understood, successive registration seems to be equivalent to simultaneous registration.

The pattern of an optic array was said to be a 'stimulus' for an eye. For animals with fixed eyes on either side of the head this assertion is plausible. Each eye registers its own hemisphere. For animals with mobile eyes at the front of the head, however, the assertion is not so obvious. The total pattern of the spherical array is then, strictly speaking, a *potential* stimulus as distinguished from the *momentarily effective* stimulus which exists for a single eye-posture. The pattern of a single entering sheaf of rays is only a sample of the whole pattern projected to a point in space. But the whole pattern may and will be picked up if the animal looks around. In an even more abstract sense the whole flux of focusable light pervading the medium is a potential stimulus which can be sampled at various station-points in the medium, although it must be explored by locomotor rather than by oculomotor action. It is, of course, unorthodox to assume that the focusable light converging to an eye is something to which the word 'stimulus' should be applied. We have generally believed that only the focused light constituting the retinal image excites receptors. But this physiological conception of the stimulus has been a source of paradox and confusion in psychology. In fact, it does not apply. The image is a stimulus for the retinal mosaic, to be sure, but not for an eye. Focusable light is the stimulus for an eye, which responds first by focusing it. The image is no more than a *response-produced* stimulus. A retinal image is not a thing with definite boundaries in any case. The retina continually moves behind it, with both large and small excursions, so as to bring the fovea to different bits of detail. A continually new sample of the pattern of the world gets projected on the retina. Images are the necessary means by which an animal exploits the capacity of light to project reflecting surfaces to any given point in a transmitting medium. The light converging to such a point is the external, stable, and relatively permanent stimulus for an eye.

(5) *The presence or absence of a solid environmental surface in any part of the optic array is specified by the textured or textureless character, respectively, of that part. In other words, textured light indicates the differentially reflecting structure of a solid surface, which stops locomotion. Untextured or homogeneous light indicates an unobstructed medium of water or air, into which locomotion can proceed indefinitely* (Gibson & Dibble, 1952). *The presence of a solid* object *in the environment is specified by a textured cone in the optic array with a*

distinct boundary, and the absence of an object is guaranteed by the absence of such a stimulus (Gibson, 1951).

This assumption, it may be noted, contradicts the classical distinction between sensory and perceptual processes. The present tacit agreement as to where the responsibility of the visual physiologist ends and the responsibility of the visual psychologist begins is thereby destroyed.

(6) *In so far as the physical solids of the environment are rigid and permanent (as contrasted with fluid and movable) the pattern of the optic array at any point in the medium is unique and permanent. That is to say, there is, and will be, one and only one pattern for each station point. It is also assumed that no considerable section of a natural environment is a duplicate of another—that models or copies of environing spaces are created only by man. An important consequence of this assumption is that the* difference *between the patterns of the optic arrays at any two station points is a unique difference which depends only on the difference between the points* (Nicod, 1930). *We may call it a unique transformation.*

(7) *An animal which moves passes through a continuous series of station points. Each eye is therefore presented with a continuous family of transformations, and this family is unique to the particular path of locomotion. With an animal which does* not *change its position, the eyes are presented with what may be called a continuous non-transformation of the optic array, that is, a static pattern.*

The classical geometry of perspective will describe the static pattern of the optic array, that is, the projection of an environment to a stationary point. The geometry of 'motion perspective' (Gibson, Olum & Rosenblatt, 1955) will describe the 'flow pattern' of the optic array, that is, the projection of an environment to a moving point. The geometry of motion perspective has been exemplified by an analysis of aerial flight with reference to a plane surface. It is interesting to note that the flow pattern is independent of the static pattern that carries it; the same flow pattern may be embodied in any number of static patterns.

(8) *An eye is a device which registers the flow pattern of an optic array as well as the static pattern of an array. Conversely, such a family of continuous transformations is a stimulus for an eye. There are quite specific forms of continuous transformation, and the visual system can probably discriminate among them* (Gibson & Gibson, 1957). *This mode of optical stimulation is an invariable accompaniment of locomotor behaviour and it therefore provides 'feedback' stimulation for the control and guidance of locomotor behaviour. It might be called* visual kinaesthesis.

The last assumption asserts something like an unrecognized sense modality. Visual kinaesthesis is, of course, supplementary to the recognized mode of proprioceptive kinaesthesis. It differs, however, in several ways. First, it seems to provide information about movements of the animal relative to the environment, not about movements of parts of the body relative to other parts, as the muscle-sense does. Secondly, it seems to provide information about displacements rather than information about accelerations and gravitational forces, as the vestibular sense does. Thirdly, the displacements registered have reference to the stable solid surfaces of the environment; displacements with reference to the medium of air or water, in the case of flying or swimming animals, are given only by proprioceptive kinaesthesis. Kinaesthesis has long been defined as the sense of bodily motion. It was described by Sherrington fifty years ago (1906). It depends on the sensitivity of receptors in the muscles and joints to compression, on the sensitivity of a statocyst to force, and also on the sensitivity of the skin to deformation. Visual kinaesthesis

depends on the sensitivity of a retinal mosaic to an overall change of pattern. To suggest that the eye can serve as a kinaesthetic organ is somewhat upsetting; for one thing it violates our notion of the eye as an exteroceptor in Sherrington's classification of exteroceptors, proprioceptors, and interoceptors. There is some inadequacy, however, in all attempts at anatomical classification of fixed senses. Kinaesthesis, as usually understood, is actually several different forms of sensitivity to several different kinds of motion: of the muscles, the joints, the endolymph, and the skin. Why not admit optical motion? The fact is that animals seem to pick up stimulus information from the sea of energy around them through multiple channels and in various ways. Seemingly redundant information is not neglected; on the contrary, many supplements and concomitants and combinations of energy seem to be received at the same time. And the transformations in the field of view are very useful information.

Consider the case of a fish maintaining its position over the bottom of a flowing brook, or the case of a bird hovering over the same bit of land in a wind. It must move relative to the medium in order to be motionless relative to the solid substratum. For this, obviously, it must rely on vision. It has to maintain an unchanging pattern of the optic array. Transformation of the array is an index of motion relative to the substratum, and a completely reliable one. By visual kinaesthesis the animal can co-ordinate its behaviour to the environment while at the same time by muscular kinaesthesis it can co-ordinate its locomotion to the medium.

Is there evidence to show that the ocular mechanism in animals can actually register the flow pattern of an optic array, as assumed? Animals, including insects, are said to be very sensitive to 'motion', and there is observational evidence from zoology and ethology (Tinbergen, 1951) suggesting that animals make specific responses to specific kinds of moving objects in the environment. If they are sensitive to objective motion in the environment it is likely that they are also sensitive to subjective movement in the environment (Gibson, 1954), but the question has not been specifically studied. There is, of course, plenty of evidence to show that animals have pattern vision in varying degrees, and experiments on form discrimination make a large literature in comparative psychology. Whether they have *change-of-pattern* vision is another question, and this has scarcely been investigated. But it is very likely that they do. The compound eyes of insects, particularly those with a relatively coarse mosaic, may prove to have greater acuity for differences between transformations than they do for differences between forms. There is evidence to prove that human observers are very sensitive to optical motions. They can detect motion parallax in the field of view with very great precision (Graham, 1951), and they can utilize the shadow transformations on a translucent screen for accurate judgments of changing depth (Gibson & Gibson, 1957).

III. Formulae for a theory of visual control of locomotion

We are now prepared to apply the postulates to the problem of how an animal gets about in a stable environment, approaching goals, avoiding obstacles, and even responding to moving objects such as a prey or a predator. By these assumptions, transformations of the total optic array to a moving point serve for the control of locomotion relative to the total environment. Transformations of a smaller bounded cone of the optic array (a closed contour with internal texture in the visual field of the animal) serve for the control of

locomotion relative to an object in the environment. A formula will be given for each of the following aspects of locomotor behaviour: beginning forward locomotion, ceasing locomotion, reversing locomotion; steering toward a specific place or object; approaching without collision; avoiding obstacles; pursuit of a moving object; and avoiding a moving object. The formulae will be in verbal rather than mathematical form, but they are capable of the latter kind of statement.

(i) *Starting, stopping, and backing-up*

The flow pattern of the optic array during forward displacement vanishes during non-displacement and is reversed during backward displacement (seventh assumption). A human observer who attends to his visual sensations can note a sort of 'expansion' or 'contraction' of the patchwork of colours in his frontal visual field as he moves forward or backward. Actually there is a centrifugal flow of the structure of the optic array from a pole in the direction of displacement, the flow being graded in proportion to the nearness of the corresponding surface (Gibson *et al.* 1955). (This graded centrifugal flow of an optic array from an environment of solids has the interesting property of being capable of continuing indefinitely without the consequence that the pattern as a whole becomes larger. Elements or forms of the pattern are magnified as the corresponding objects get nearer, but new elements emerge to take their place in the array. The structure of such an array consists of forms within forms to an unlimited density. The reason for this fact is that the environment of animals is itself structured at any level of magnitude; it is composed not only of mountains and trees but also of crystals and cells.) There is a corresponding centripetal flow toward a pole in the direction opposite the displacement. These flow patterns depend on the configuration of the reflecting surfaces in the environment but they are independent of any *particular* configuration. To begin locomotion, therefore, is so to contract the muscles as to make the forward optic array flow outward. To stop locomotion is to make the flow cease. To reverse locomotion is to make it flow inward. To speed up locomotion is to make the rate of flow increase and to slow down is to make it decrease. An animal who is behaving in these ways is optically stimulated in the corresponding ways, or, equally, an animal who so acts as to obtain these kinds of optical stimulation is behaving in the corresponding ways.

(ii) *Steering and aiming*

The centre of the flow pattern during forward movement of the animal is the direction of movement. More exactly, the part of the structure of the array from which the flow radiates corresponds to that part of the solid environment toward which he is moving. If the direction of his movement changes, the centre of flow shifts across the array, that is, the flow becomes centred on another element of the array corresponding to another part of the solid environment. The animal can thus, as we would say, 'see where he is going'. The act of turning or steering is, therefore, a visual as well as a muscular event. To turn in a certain direction is to shift the centre of flow in that direction relative to the fixed structure of the optic array. The amount of turn is exactly correlated with the angular degree of shift. The behaviour of aiming at a goal object can now be specified (although the properties of a figure in the field of view which arouse this behaviour have not yet been described). To aim locomotion at an object is to keep the centre of flow of the optic array as close as possible to the form which the object projects.

(iii) *Approaching without collision*

Approach to a solid surface is specified by a centrifugal flow of the texture of the optic array. Approach to an object is specified by a magnification of the closed contour in the array corresponding to the edges of the object. A *uniform* rate of approach is accompanied by an *accelerated* rate of magnification. At the theoretical point where the eye touches the object the latter will intercept a visual angle of 180°; the magnification reaches an explosive rate in the last moments before contact. This accelerated expansion in the field of view specifies imminent collision, and it is unquestionably an effective stimulus for behaviour in animals with well-developed visual systems. In man, it produces eye blinking and aversive movements of the head, even when the stimulus is a harmless magnification of a shadow on a translucent screen. At lesser intensities this 'looming motion', as it might be called, presumably yields lesser degrees of aversion, or a slowing down of approach. The fact is that animals need to make contact without collision with many solid objects of their environment: food objects, sex objects, and the landing surfaces on which insects and birds alight (not to mention helicopter pilots). Locomotor action must be a balance between approach and aversion. The governing stimulation must be a balance between flow and non-flow of the optic array. The formula is as follows: contact without collision is achieved by so moving as to cancel the centrifugal flow of the optic array at the moment when the contour of the object or the texture of the surface reaches that angular magnification at which contact is made.

(iv) *Steering among obstacles*

Fish, birds and arboreal animals live in a cluttered environment, that is, one whose open spaces are encroached upon by solid surfaces. Nevertheless, they swim, fly, or take a course among these obstacles with great precision. Many species do so only when the medium is illuminated, that is, they depend on vision. The obstacles are specified in the optic array by contours with internal texture; the open spaces are specified by the areas between such contours. The background areas may be untextured homogeneous colour, like the sky, or densely textured surface colour, like the earth. The question is, what governs the taking of a course during such locomotion? Symmetrical magnification of a textured contour specifies a collision course toward an obstacle, but a skewed magnification, where the centre of flow is outside the textured contour, specifies a non-collision course. So long as an animal keeps the focus of centrifugal flow outside the textured contours and within a homogeneous or densely textured area, he will not collide with a solid object. This formula, it may be noted, is simply the reciprocal of that for aiming at a goal object. The properties of a textured contour by which is specified either an obstacle on the one hand or a goal object on the other have not yet been considered.

(v) *Pursuit and flight*

Predatory animals pursue and preyed-upon animals flee. These biologically complementary forms of behaviour must both be controlled by transformations of optical pattern. Approach to a stationary goal object has been described. A fugitive goal object is specified when the optic array as a whole flows from a centre, but when a textured contour within it does not expand. The object is then moving away and being pursued. Absolute angular magnification of the contour means catching up and minification of the contour means falling behind. Hence, the rule by which a big fish can catch a little fish is

simple: amximize its optical size in the field of view. From the point of view of the prey, the expansion of a textured contour in the optic array means the approach of something. This in itself may touch off the reaction of flight if it comes within the field of view. Preyed-upon animals need eyes which can register this abstract optical event in the rearward as well as the frontward direction, and very generally they seem to have developed eyes with a panoramic binocular field (Walls, 1942). What exactly *is* flight? The formula is again simple. It is so moving as to minimize the optical size of the expanding contour. This will necessarily involve a centripetal flow of that hemisphere of the optic array which contains it and a centrifugal flow of the opposite hemisphere. If the optical size of the contour decreases, the animal is getting away; if it increases despite the contraction of the surrounding field, the animal is being overtaken. There is one other geometrical possibility of stimulation: contraction of a contour in an otherwise static array. This means something *going* away. Probably it does not touch off the reaction of flight as expansion of a contour does in preyed-upon animals. This hypothesis needs empirical testing.

(vi) *The complexities of terrestrial locomotion*

For simplicity, the formulae of the theory have referred, so far, to locomotion in a medium of water or air. They are sufficient to show the possibilities of a circular control or 'feedback' type of stimulus-response analysis for understanding the locomotor behaviour of fish and birds. There are, however, terrestrial animals. The surface of the earth is a substratum. Locomotion can occur, mechanically speaking, in two possible ways: by exerting force against the medium and avoiding contact with the substratum, or by exerting force against the substratum and keeping in more or less continuous contact with it. The latter type, terrestrial locomotion, seems to have developed later in evolution. Terrestrial animals utilize a great variety of mechanical procedures—crawling, gliding, hopping, walking, running and the like, but they all push against the ground in one manner or another. For these animals, touch is available to play some part in controlling locomotion—actually *change* of contact stimulation, or tactual motion. There is also classical muscle-joint kinaesthesis which, in these animals, is never discrepant with visual kinaesthesis as it may be for the fish or the bird in a flowing medium. Since these modes of controlling stimulation run parallel in terrestrial man, it has been easy for the human psychologist to overlook the visual mode. Only since man began to fly aircraft has the visual component in the guidance of locomotion forced itself on the attention of psychologists. And yet it was always there to be observed. Consider the perception of 'footing' in a man who is making his way over rough terrain. Or, better, consider the oculomotor co-ordinations of a cat, or of any animal which runs with 'due regard' to the footing. The visual aspect of the performance is just as precise as the muscular aspect. The animal has only the information supplied by the optical array to a moving point (or a pair of them) and yet it adjusts its behaviour to the convexities and concavities and other physical properties of the surface ahead of it. This consideration reminds one that animals respond to invariants as well as changes of stimulation—to the permanent properties of the environment as well as to their own motions in it. We should return to the problem of the exteroceptive function of vision and note whether the problem is reformulated after first considering its proprioceptive function. Animals are capable not only of locomotion in, but also of orientation to, their environment, that is, of object perception and even what psychologists have called 'space' perception.

IV. Visual orientation in animals

Animals are not continually on the move; they sometimes remain at rest. Often the only movements that can then be observed are eye-movements. This means that the resting animal is exploring the optic array even if he is not exploring the environment. It is tempting to believe that the animal is still making some kind of implicit responses to the solid surfaces around him even if he is not explicitly poking his nose into things. He is certainly oriented to gravity and to the substratum when he maintains a constant posture. He is probably also oriented to the distant objects of his environment when he remains in the same place. In the fifth assumption, we may indeed find a basis for the latter kind of orientation. It seems to be possible for him to identify and react to the objects of his world by virtue of the optical properties of the light reflected from them to his eye. One might even wish to suggest that the animal, in his way, *perceives* the objects of his world.

Animals make different kinds of locomotor reactions to different objects. They approach food or shelter, they avoid obstacles, they pursue prey and they flee the predator. These are discriminative reactions and they require a different kind of stimulus-response theory than do the control reactions heretofore considered. We must now consider actions which are specific to those features of the optic array which do *not* change during locomotion rather than those which do. Such features of stimulation are not response produced and the responses are not circular. In such behaviour the S-R linkage is between permanent entities of the environment and acts which are appropriate to them. The distinction between an S-R theory of *control* reactions and an S-R theory of *identifying* reactions is important for behaviour theory. It is true that an automaton can be designed which will aim at, approach, and pursue a pre-set target (as witness military missiles) and that no automaton has yet been designed which will recognize targets appropriate to its own needs (apart from its designer's) and act accordingly. But it would be wrong to categorize the first kind of reactions as *automatic* and the second kind as *voluntary*. This dichotomy is as pernicious as the one between sensory and perceptual processes. The true distinction is probably that between the properties of stimulation which vary over time and those which do not.

(i) *Object perception*

By the fifth assumption, material objects, substances in the solid state, are specified by the textures and contours of the optic array, and they stop locomotion. However, a solid object, by virtue of its chemistry and biochemistry, may constitute food for an animal. Or it may constitute a mate, or young, or shelter, or an obstacle, or an enemy. These higher-order properties of solids are specified by higher-order properties of the light in the optic array. For example, there is internal pattern as well as texture in the light reflected from an object and there is shape as well as contour. There is a frequency distribution in such light which specifies in a complicated way what men call 'colour'. Moreover, if the object is alive and moves there will be deformations of its contour and internal pattern which specify it as animate, and which may further specify it as one or another kind of animate thing. In so far as an animal can discriminate these variables of the optic array, he can discriminate the properties of objects which render them not only bump-into-able and walk-on-able but also mate-with-able, or get-underneath-able, or edible, or likely to cause pain. And to this extent, he can identify the significant classes of objects

in his environment, that is, he can respond to them differently as the objects themselves differ in biophysical ways (Gibson & Gibson, 1955). These are the so-called 'goal-objects' of an animal's environment, the ones which are related to the animal's needs, which induce approach or avoidance, and which man describes as having valences or meanings in the perceptual field. The hypothesis is that the values of objects as well as their solidity are specified with some degree of reliability in the optic array. Mere physical solidity has a value for locomotion and collision. Chemical and biological properties of higher order have values for nutrition, reproduction, and survival. The animal's task is to detect these properties at a distance.

Contemporary behaviour theorists either take for granted that objects constitute stimuli ('stimulus objects') and thus beg the question of how objects are responded to, or else they assume vaguely that pattern stimuli are 'cues' for behaviour and thus slip around the problem of object perception. (The latter point is illustrated by comparing the use of 'cue' by a student of behaviour such as Miller—Miller & Dollard, 1941—who takes the concept very lightly, and a student of perception such as Brunswik, 1956, who takes the concept very seriously.) But this will not do. During locomotion, the patterns at the receptor surfaces of an animal are necessarily in flux, the objects being projected as changing patterns rather than constant patterns. If we wish to believe that an animal can respond to a constant object in the course of his behaviour we should first resolve this apparent contradiction. There are two alternatives. We can assume that the animal somehow constructs a perceptually constant object from the kaleidoscopic field of sensations: this is the starting point for classical theories of perception. Or we can assume that the animal is sensitive to the mathematically invariant properties in the stimulus flux which correspond to the physically constant object: this is the alternative which seems more promising for the future. During the locomotor transformations of the pattern of the optic array the property of colour remains, of course, invariant. During these perspective transformations many geometrical properties such as the kind of texture, or the angularity or curvature of the pattern also remain invariant. There are plenty of invariant variables in the flux of optical stimulation on the basis of which an animal can identify classes of permanent objects. The empirical question is whether animals are, or come to be, sensitive to them.

(ii) *Orientation to the visible environment*

Animals seem not only to recognize constant objects, but to orient themselves to a constant environment, that is, to any permanent arrangement of solid surfaces. The homing behaviour of many species suggests that the animal is able, on familiar ground, to find its way from any one place to any other significant place; territorial behaviour also suggests that he can respond, in some sense, to the habitat as a whole; and exploratory behaviour suggests a kind of striving to extend the boundaries of this whole. How can this complicated behaviour be supposed to depend on vision? Consider the pattern of an optic array to a stationary position. For land animals the lower hemisphere, roughly, will be textured and patterned by the surfaces and objects of the earth. But not only solidity is specified in the array; the slants and slopes and the facings of this arrangement of surfaces are specified by *gradients* of the optical texture (Gibson, 1950, chap. 6). The edges are specified by *steps* of the texture variables, and convexities or concavities are specified by *changes* of gradient. In short, the lay of the land, the jumping-off places, the

interspaces, barriers and obstacles, as well as the level stretches, are given by the geometry of the optic array. Depending on the locomotor capacities of the animal, this terrain provides definite possibilities or impossibilities for crawling, walking, climbing and the like. And if the animal can discriminate the textural variables it can discriminate among potential paths for locomotion. A potential path is a stretch of surface extending away from the animal which affords the kind of locomotion for which the animal is equipped. A barrier or obstacle is a surface which does not afford locomotion. There are transitional cases between path and barrier, to be sure, and a cliff or vertical edge is a special type of barrier, but the definitions will serve well enough for theoretical purposes. (The responses of rats to a 'visual cliff' have recently been studied under laboratory conditions—Walk, Gibson & Tighe, 1957—and they show the kind of aversion predicted by this theory.) A terrestrial animal is always encircled, then, by a radiating set of paths or barriers. Each angular sector of the optic array specifies the possibility of locomotion in that direction. A level path of unlimited distance is given by a gradual increase of textural density up the array to the horizon. A barrier is given by a region over which density remains constant, and its distance by the point at which the change in the density gradient begins. A falling-off place is given by an abrupt increase in density upward in the array. A margin between land and water (a barrier for a non-swimmer) is given by a change in the kind and rigidity of texture together with a change in colour. Confinement consists of barriers in all directions; freedom consists of paths in all directions. To the extent that the animal can respond to these variables of the encircling optic array, we may now conclude, he has locomotor orientation to his environment. He perceives the possibilities of locomotion surrounding him. *And this is probably what should be meant by asserting that an animal has visual space perception.*

(iii) *Orientation to the environment outside the range of vision*

The visible environment may be limited by enclosing surfaces such as those of a laboratory maze, and even in an open medium it is limited by the acuity of the eye, the clearness of the medium, and the level of the illumination. An animal, nevertheless, can learn in some degree to go to places and objects outside of these limits (Rabaud, 1928; Thorpe, 1956) just as a man can learn to find his way about a house, the streets of a city, or the highways of a state. How is this behaviour to be accounted for, and in what sense is it visually determined? It is not explained by the unchanging pattern of an optic array to a single position. Consider, however, the transformations of pattern of an optic array to a moving position, as described in the seventh and eighth postulates. We have assumed that a family of continuous transformations is itself an optical stimulus, and that the properties which remain invariant under transformations also remain stimuli during such flux. The former specifies the particular path of locomotion taken; the latter specify the permanent arrangement of surfaces. An animal that explores its environment in the course of time moves from each of many places to each of many other places and, just as each place corresponds to a unique pattern, so each movement corresponds to a unique transformation. The visual flux of a locomotor animal, as it goes about its business, consists of a sequence of transformation families interspersed with periods of non-transformation. Since these families each begin where the preceding one ended, they are linked. The entire life of the animal is co-extensive, in fact, with a single grand family of optical transformations which specify the history of its travels and explorations. This fact should

be taken into consideration in any theory of maze learning or, more generally, of how an animal learns its way about. The successive patterns of the optic array are not a series of discrete pictures like a series of nonsense syllables which have to be associated; they are parts of a continuous transformation with a temporal pattern of its own. The nature of this temporal pattern—the direction of the change of the momentary patterns—is probably something which an animal can sense and, if so, we do not have to postulate the conditioning of a series of responses to a series of discrete stimuli as Hull (1952) does in his theory of the goal-gradient and the habit-family. Hull chops a temporal pattern into a set of static patterns and is then faced with the problem of integration. The integration of behaviour is partly explainable in terms of the pre-existing integrity of stimulation, and by the hypothesis that an animal learns to respond to it. If maze learning is to be explained in S-R terms we must recognize that the stimuli *to which responses are made* change as well as the responses themselves (Gibson & Gibson, 1955). In the course of learning, a whole temporal pattern may become the effective stimulus for a single integrated act.

The controversy over 'response learning' and 'place learning' has now reached major proportions in psychology (Woodworth & Schlosberg, 1954, chap. 21; Thorpe, 1956, chap. 5), but neither side has troubled to analyse the optical stimulation. Hull's theory of maze learning asserts that it is the acquiring of a sequence of movements. Tolman's theory asserts that it is the acquiring of a cognitive map of the maze (Tolman, 1948). If Hull's theory could profit by utilizing the notion of continuous transformation, Tolman's theory could profit by the notion of invariant variables of stimulation under transformation. The perceiving of the possibilities of locomotion surrounding an animal does not, as Tolman assumes, require a theory of sign learning or the interpretation of sensory cues. Space perception can be explained with greater parsimony by kinds of stimulation hitherto neglected in sensory physiology. Knowing the possibilities of locomotion *outside* the limits of momentary vision, that is to say the cognitive mapping of the extended environment, can be explained in part by the recurrent, constant, or invariant properties of such stimulation which are discovered during exploratory behaviour.

(iv) *The learning of visual orientation by exploration*

Animals learn to get about the environment and they also learn *about* the environment. The first kind of learning is called behavioural and the second cognitive, but nothing prevents us from making both assertions. If an animal goes to where something is we can infer that he knows where it is. In either case the experimental evidence suggests that the learning depends on *exploratory* activity. Artificial environments such as mazes have provided much of the evidence, but the conclusion holds for natural environments as well, the evidence coming from work on homing and foraging (Thorpe, 1956). No other kind of stimulation is as informative about the environment as light when it is many-times reflected in a medium. The best means of exploring an environment is (*a*) to point the eyes in various directions, and (*b*) to plant them at various station points. It seems probable that eyes are sensitive to the transitions between visual patterns as well as the patterns themselves. Certainly locomotion is guided thus. A transformation may be more significant than a form, in fact, since it carries all the information that a form does plus that carried by the kind of transition. There are families and superfamilies of such information in focusable light. The more of it an animal can react to the more efficiently and

quickly he can go from place to place in his environment, and the more he can learn what the arrangement of his environment is. This kind of learning is by no means the only kind, to be sure, but it is an important one.

References

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments.* Berkeley: University of California Press.

Fraenkel, G. S. & Gunn, D. L. (1940). *The Orientation of Animals.* London: Oxford University Press.

Gibson, J. J. (1950). *The Perception of the Visual World.* Boston: Houghton-Mifflin.

Gibson, J. J. (1951). What is a form? *Psychol. Rev.* **58**, 403–12.

Gibson, J. J. (1954). The visual perception of objective motion and subjective movement. *Psychol. Rev.* **61**, 304–14.

Gibson, J. J. & Dibble, F. N. (1952). Exploratory experiments on the stimulus conditions for the perception of a visual surface. *J. Exp. Psychol.* **43**, 414–19.

Gibson, J. J. & Gibson, E. J. (1955). Perceptual learning: differentiation or enrichment? *Psychol. Rev.* **62**, 32–41.

Gibson, J. J. & Gibson, E. J. (1957). Continuous perspective transformations and the perception of rigid motion. *J. Exp. Psychol.* **54**, 129–38.

Gibson, J. J., Olum, P. & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *Amer. J. Psychol.* **68**, 372–5.

Graham, C. H. (1951). Visual perception. Chap. 23 in S. S. Stevens (Ed.), *Handbook of Experimental Psychology.* New York: Wiley.

Hull, C. L. (1938). The goal-gradient hypothesis applied to some 'fieldforce' problems in the behavior of young children. *Psychol. Rev.* **45**, 271–99.

Hull, C. L. (1952). *A Behavior System.* New Haven: Yale University Press.

Lewin, K. (1933). Environmental forces. Chap. 14 in C. Murchison (Ed.), *A Handbook of Child Psychology.* Worcester: Clark University Press.

Lewin, K. (1935). *A Dynamic Theory of Personality.* New York: McGraw-Hill.

Miller, N. E. & Dollard, J. (1941). *Social Learning and Imitation.* New Haven: Yale University Press.

Nicod, J. (1930). *Foundations of Geometry and Induction.* London: Kegan Paul, Trench, Trubner and Co.

Rabaud, E. (1928). *How Animals Find their Way About.* New York: Harcourt, Brace.

Sherrington, C. S. (1906). On the proprioceptive system especially in its reflex aspect. *Brain*, **29**, 467–82.

Thorpe, W. H. (1956). *Learning and Instinct in Animals.* London: Methuen.

Tinbergen, N. (1951). *The Study of Instinct.* London: Oxford University Press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208.

Walk, R. D., Gibson, E. J. & Tighe, T. J. (1957). Behavior of light- and dark-reared rats on a visual cliff. *Science*, **126**, 80–1.

Walls, G. L. (1942). *The Vertebrate Eye.* Bloomfield Hills, Mich.: Cranbrook Institute of Science.

Woodworth, R. S. & Schlosberg, H. (1954). *Experimental Psychology.* London: Methuen.

(*Manuscript received* 28 *June* 1957)

EXPERIMENTS ON 'NEOPHOBIA' IN WILD AND LABORATORY RATS

By S. A. BARNETT

Department of Zoology, University of Glasgow

All rats (*Rattus norvegicus*), wild or tame, have a well-marked tendency to explore their surroundings. Wild rats, but not tame ones, also tend to avoid unfamiliar objects in a previously explored, and therefore familiar, area. The effect of environmental change on feeding was studied in three groups of wild rats. Rats of Group 1 were accustomed to feeding from a wire basket at the back of the cage; for them the change consisted of transfer of the food to an unfamiliar tin at the front of the cage. Group 2 had food in the basket throughout, and the change consisted only in an empty, unfamiliar tin being put in the front of the cage. The rats of Group 3 fed from a basket in the front of the cage; the change for them was the placing of an unfamiliar empty tin between the back of the cage, where they slept, and the food. In all three groups feeding was interrupted by the change. In the first group, three of the five rats stopped eating completely for one or more days; in the other two groups, food consumption was merely reduced on the first day after the change. The check in eating was not due to exploration: the new object was at first completely avoided by all the rats. Tame albino and hooded rats, subjected to the same procedure as the wild rats of Group 1, investigated the tin and the food in it at once; their food intake during the day after the change was unaffected. The first approach of the hooded rats to the new object was, however, more hesitant than that of the albinos. Investigation of a new object, or new conditions, can sometimes lead to an interruption of feeding in tame rats, but this has nothing in common with avoidance behaviour in wild rats.

I. Introduction

During the movements of wild rats about their living space, two apparently opposed types of response may be observed. The first is exploration: this is especially clearly seen when a new area becomes accessible (Barnett & Spencer, 1951), or when rats are experimentally introduced into an unfamiliar cage or enclosure (Barnett, 1958*a*); this sort of behaviour is highly developed in tame laboratory rats as well as in wild ones. The second type of response is the avoidance of unfamiliar objects in familiar surroundings (Chitty & Shorten, 1946); this 'neophobic' behaviour, as will be further shown below, does not occur in typical laboratory-bred rats. The experiments described below were designed to clarify the relation between exploration and avoidance, with special reference to the differences between wild and tame rats. A general discussion of exploratory and allied behaviour will appear seperately (Barnett, 1958*b*).

II. Experimental

The rats used were healthy, adult, virgin females, born in the laboratory. The wild rats were aged 63–131 days at the beginning of the experiments in which they were used; their initial weights were 115–205 g. The tame rats were of two types, neither rigorously inbred: they were (i) albinos, aged 66 days and weighing 132–157 g., and (ii) hooded rats, aged about 70 days and weighing 150–190 g. Unlike the wild rats, they could, as usual, be readily handled without special precautions. Each rat was kept alone in a wire cage measuring 36 by 25 by 16 cm. high, and had been in the cage at least 7 days before observations began. The floor of each cage was raised; 5·5 cm. below it was a tray containing peat moss covered with thick paper; food residues fell on this paper and were collected. The weight of food eaten by each rat was recorded daily. Both before and during the experiments the only food supplied

was diet 41 (Bruce & Parkes, 1949). Excess water was always available, supplied from an inverted bottle on top of the cage.

To illustrate the effect of changed conditions on food consumption, three types of experiment were carried out. In the first, the food was initially supplied in the usual manner, in a wire basket suspended from the cage roof; this basket, and the water source, were at the back of the cage, that is, the part of the cage near the wall and farthest from people (Fig. 1); this was where the rat usually slept. When the rat was fully accustomed to this arrangement, its daily food consumption was recorded for 4 days. On the fourth day, a tin was put in the front of the cage, fixed to the wire, and food was supplied in this; the basket in which food had hitherto been supplied was left empty. This type of experiment was carried out with five wild rats, five albinos, and five hooded rats.

In the second type of experiment the procedure was the same except that (*a*) food was supplied in the basket throughout, and the tin was left empty, and (*b*) the change was made after only 3 days. In this case, then, the only change produced was the introduction of an unfamiliar object. This type of experiment was carried out with five wild rats.

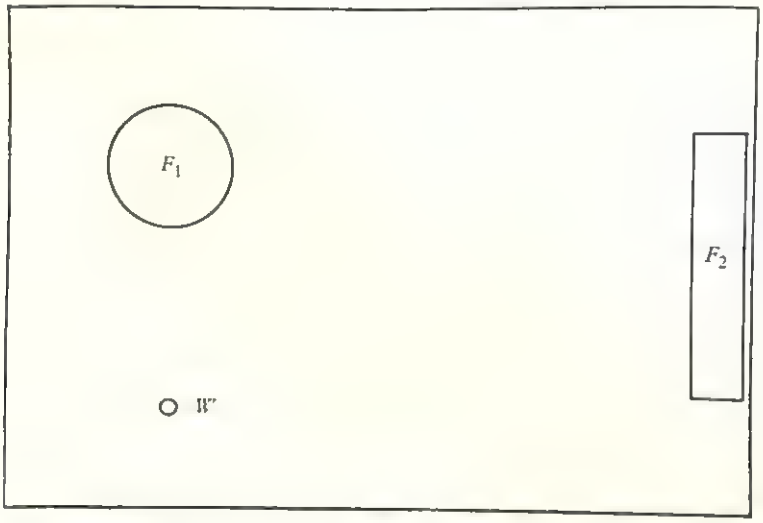


Fig. 1. Arrangement of cage for experiments of types 1 and 2. F_1 = food basket; W = water source; F_2 = food tin. Back of cage on left.

In the third type of experiment, the food basket was at the front of the cage instead of the back, while the water source was still at the back. Consequently, in order to feed, the rats had to come forward from where they usually slept. After 3 days, an empty tin was put in each cage (Fig. 2), so that this unfamiliar object had to be passed when the rat approached the food. Food was supplied in the basket throughout, never in the tin. This type of experiment was carried out on six wild rats.

The room in which the experiments were done contained many other rats. Its temperature was kept between 19 and 23° C., and it was continuously ventilated by an extractor fan. There is no reason to think that there was any substantial variation in the olfactory stimuli to which the rats were subjected during the experiments.

III. Results

The most drastic alteration of conditions was that imposed in the experiments of the first type, in which the food was transferred from a basket at the back of the cage to a completely unfamiliar tin at the front. In the first phase of this experiment, before the changeover was made, the wild rats always began to eat within 3 hr. of the replenishing of the food baskets (at about 11.00 hr. G.M.T.); this was shown by the presence of kibbled fragments below the cage floor. After the change on the fourth day, no rat had fed at all by 18.00 hr., that is, 7 hr. after the new arrangement had been made. All droppings

were at the back of the cage, showing that the rats had remained there and that none had investigated the new object or had made any general search for food.

The daily intake of food is shown in Fig. 3; in this case, in view of the variation between individual rats, the record is given separately for each. Despite the altered behaviour of all rats during the first 7 hr., two of the rats (nos. 4 and 5) showed only a slight decline in food consumption on the fifth day. The other three ate no food at all on that day; of these, no. 3 began to eat on the seventh day, no. 1 on the eighth day, and no. 2 on the ninth. The position of kibbled remnants below the cage floor showed that all eating was done, by all the rats, at the back of the cage. Further, once they had begun to feed, all the rats developed the habit of removing the food cubes from the tin to the back of the

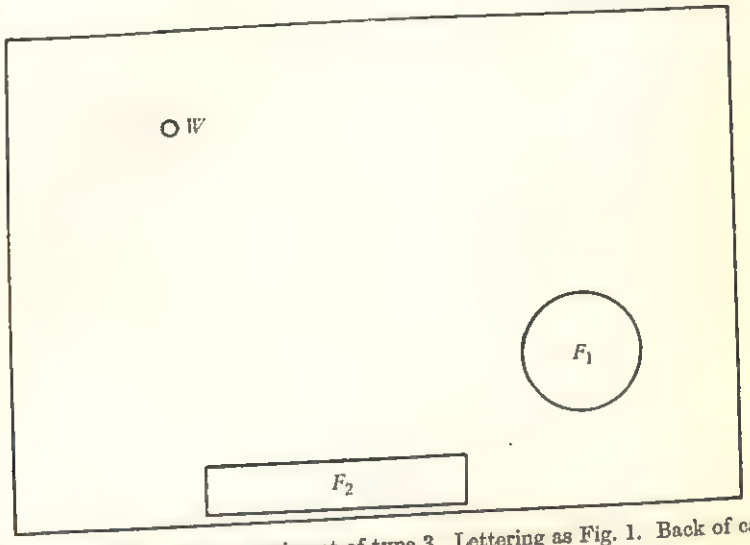


Fig. 2. Arrangement of cage for experiment of type 3. Lettering as Fig. 1. Back of cage on left.

cage. On the tenth day, all food had been removed from all tins, except for four cubes left in the tin by rat no. 5. At the end of the experiment, proper food was once more put in the basket, as well as in the tin; but the rats continued to remove the cubes from the tin and to collect them at the back of the cage.

The effects of the less drastic changes in the experiments of the second and third types are shown in Figs. 4 and 5. In the experiments of the second type, all the rats showed some reduction in food consumption during the 24 hr. after the introduction of the tin, but all ate some food. The mean consumption on day 4 is significantly different from the mean consumption on day 3 ($P \leqslant 0{\cdot}05$). Judged by the absence of droppings in the front of the cage, the tin was entirely avoided during this 24 hr. period: the rats evidently remained at the back of the cage and refrained from exploration. In subsequent days, droppings were seen in front of the cage; but no detailed observations were made on the stages by which the 'shyness' of the new object was overcome.

In the experiments of the third type, one of the rats ate the same weight of food on the fourth day as on the third, but the rest all ate less; the mean consumption of food on day 4 was however not significantly different from that on day 3 (Fig. 5). Nevertheless, as in the other experiments, there was a delay in visiting the front of the cage: there was no evidence of feeding or of exploration during the first 5 hr. after the change had been

made. In this situation the rats had to go to the front to feed, since the construction of the basket precluded carriage of food cubes to the back.

Finally, the response of the tame rats, to the first type of experimental procedure, must be described. Within 4 min. from the time when the unfamiliar tins, containing food, had been placed in the cage, all the albinos were investigating the tins and removing cubes; their approach was made without hesitation. Fig. 6 shows that daily food intake was not appreciably influenced by the change of conditions. One rat ate only 11 g. on

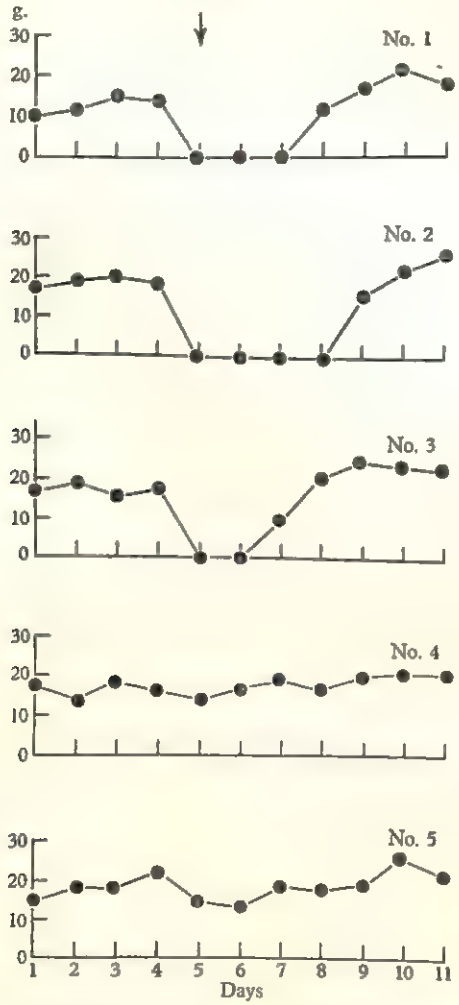


Fig. 3. Daily food consumption in g. of individual wild rats subjected to experimental procedure, type 1.

the fourth day, after eating 14, 15 and 16 g. respectively on the previous days, followed by 17 g. on the fifth day. One rat ate 3 g. more on the fourth day than on the third; the consumption of two others differed by only 1 g. from that of the day before; and that of one was unchanged. The mean food consumption on day 4 does not differ significantly from that on day 3. All rats, from the first day on which they had the opportunity, regularly removed the cubes from the tin, and scattered them about the cage.

The hooded rats also investigated the tins almost at once: four had begun to do so by

the seventh minute, but one delayed until the twentieth minute. By the fifteenth minute the first four were also removing food cubes. Fig. 7 shows that, as with the albinos, daily food intake was not materially altered. The manner of approach of these rats to the food tin differed, however, from that of the albinos: the initial movements were more tentative and hesitant; in this, they resembled those of a wild rat faced with a new object, but the hesitation was much sooner overcome. Four of these rats scattered food about the cage, but all ate usually at the back.

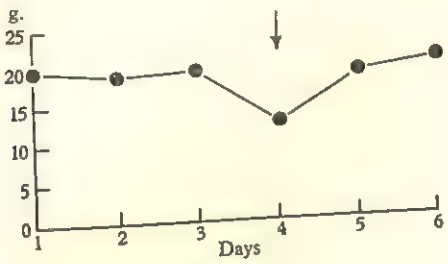


Fig. 4

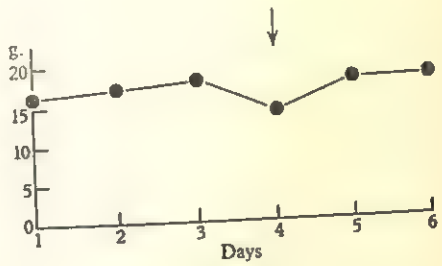


Fig. 5

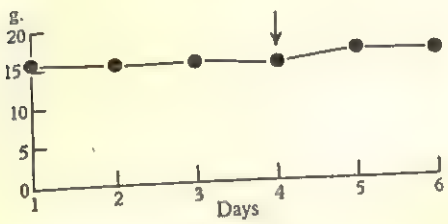


Fig. 6

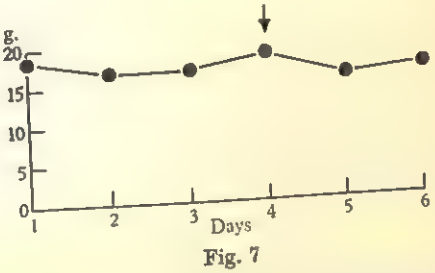


Fig. 7

Fig. 4. Mean daily food consumption in g. of five wild rats subjected to experimental procedure, type 2.
Fig. 5. Mean daily food consumption in g. of six wild rats subjected to experimental procedure, type 3.
Fig. 6. Mean daily food consumption in g. of five albino rats subjected to experimental procedure, type 1.
Fig. 7. Mean daily food consumption in g. of five hooded rats subjected to experimental procedure, type 1.

IV. Discussion

The results of all three experimental procedures show that what Chitty & Shorten (1946) call 'new object reaction' can be displayed even in a small cage. This neophobia consists of avoiding an unfamiliar object in familiar surroundings (and also, in some instances, a considerable area around it). Such a response must clearly tend temporarily to reduce food intake when accustomed food is replaced by a new food (Chitty & Shorten 1946; Chitty, 1954; Shorten, 1954), although Shorten (1954) distinguishes between new object reaction and the refusal of an unfamiliar foodstuff (cf. Barnett, 1956). The placing of a new object near or at a rat's usual feeding place has a similar effect: the behaviour involved has been extensively discussed both by Chitty and by Shorten in the papers quoted.

In the first experimental procedure the change in conditions was complex; it included the introduction of a new object, and the transfer of the food to this object. The *absence* of food from the basket which had hitherto contained it was probably unimportant (Shorten, 1954), except that it obliged the rats to get their food in an unfamiliar place.

The effect of the behaviour of three of the five wild rats, to which this procedure was applied, was drastic. The contrastingly brief interference it caused in the feeding of the other two rats of this group reflects the great individual variation which must be expected in dealing with members of a genetically heterogeneous population. This variation in behaviour has also been recorded by Chitty (1954) from observations on free-living colonies. Had food intake been recorded at shorter intervals than 24 hr., the interruption in the feeding of all five rats could have been more clearly shown.

The environmental changes imposed in the other experiments were relatively trivial. In the third type of experiment, the introduction of a new object between a rat and its food merely inhibited food consumption during the first 5 hr. of the first day, evidently by delaying the approach to the already familiar food and container. Qualitatively, this may be regarded as an example of typical new object reaction. The results of the second type of experiment were similar. Here the new object was put in the front of the cage, but the rats were still able to feed as usual at the back. Nevertheless, food intake was significantly reduced on the first day. The reduction in food eaten was in neither case due to preoccupation with exploration or investigation of the new object: on the contrary, the new object was avoided at first. Thus 'new object reaction' in these situations was clearly a result of 'fear' induced by the introduction of the unfamiliar object, presumably visually perceived, into the immediate environment of the rat. Montgomery & Monkman (1955) have shown that, in laboratory rats too, 'fear' reduces exploration.

The result of applying the first type of experiment to albino rats illustrates the complete absence of avoidance behaviour or new object reaction in any form in these animals. The response to new conditions is an immediate and vigorous exploration. There was no alteration in the daily rate of food consumption. Much the same applies to hooded rats, but in them the first approach to the new object was more wary; it is possible that this is a consequence of their superior vision since, unlike albinos, hooded rats have normal pigmentation of the iris. As already mentioned, exploration is also a conspicuous feature of the behaviour of wild rats (Barnett, 1956; see also Barnett & Spencer, 1951; Barnett, 1958*a*); but it is inhibited in the special conditions of the introduction of a new object into familiar surroundings.

The results of the experiments given above show, even without the observations of other workers, that the interference with food consumption resulting from unfamiliar stimuli is exceedingly complex. Among the processes which may be involved are (*a*) the failure to visit a place in which food has been newly placed, and (*b*) the positive avoidance of a food which, in an unfamiliar position, constitutes a 'new object'; but positive avoidance is undoubtedly the more important of these two, since, in the absence of some specific source of inhibition, every part of a rat's home range is probably visited regularly (Barnett, 1956, 1958*a*).

The use of food consumption as an index of 'new object reaction' can lead to further difficulties of interpretation. Hebb (1949, p. 194), Campbell & Sheffield (1953) and Berlyne (1955) have described how changes of environment induce exploration in laboratory rats. Chance & Mead (1955) draw attention to the fact that reduced food intake in a new situation may result either from delay in beginning to eat or from a lower rate of feeding, or both. Their experiments, which illustrate this difference, were on albino rats; consequently, in them, any interference with feeding due to novel stimuli must be attributed to the priority given by tame rats to exploration rather than eating, as first described

by Small in 1899. Chance & Mead, however, equate the effect they observed with the new object reaction described by Chitty & Shorten in wild rats. This is clearly mistaken. The experiments described above show: (i) that exploration on the one hand, and new object reaction on the other, are quite distinct; (ii) that only wild rats display new object reaction in any substantial form; and (iii) that in wild rats avoidance behaviour is by far the more important source of delayed or reduced feeding.

Thanks are due to John Keys and Margaret MacLean for technical help.

References

Barnett, S. A. (1956). Behaviour components in the feeding of wild and laboratory rats. *Behaviour*, **9**, 24–43.

Barnett, S. A. (1958*a*). An analysis of social behaviour in wild rats. *Proc. zool. Soc. Lond.* **130**, 107–52.

Barnett, S. A. (1958*b*). Exploratory behaviour. *Brit. J. Psychol.* (in the Press).

Barnett, S. A. & Spencer, M. M. (1951). Feeding, social behaviour and inter-specific competition in wild rats. *Behaviour*, **3**, 229–42.

Berlyne, D. E. (1955). The arousal and satiation of perceptual curiosity in the rat. *J. Comp. Physiol. Psychol.* **48**, 238–46.

Bruce, H. M. & Parkes, A. S. (1949). Feeding and breeding of laboratory animals. 9. A complete cubed diet for mice and rats. *J. Hyg., Camb.*, **47**, 202–8.

Campbell, B. A. & Sheffield, F. D. (1953). Relation of random activity to food deprivation. *J. Comp. Physiol. Psychol.* **46**, 320–2.

Chance, M. R. A. & Mead, A. P. (1955). Competition between feeding and investigation in the rat. *Behaviour*, **8**, 174–82.

Chitty, D. (1954). *The Control of Rats and Mice*. Vol. I. *Rats*. London: Oxford University Press.

Chitty, D. & Shorten, M. (1946). Techniques for the study of the Norway rat (*Rattus norvegicus*). *J. Mammal.* **27**, 63–78.

Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley.

Montgomery, K. C. & Monkman, J. A. (1955). The relation between fear and exploratory behavior. *J. Comp. Physiol. Psychol.* **48**, 132–6.

Shorten, M. (1954). In *The Control of Rats and Mice*. Vol. II, Chitty, D. (ed.). London: Oxford University Press.

Small, W. S. (1899). Notes on the psychic development of the young white rat. *Amer. J. Psychol.* **11**, 80–100.

(*Manuscript received* 16 *April* 1957)

THE ROLE OF THE REINFORCER AS A STIMULUS

By R. L. REID

Psychology Department, University of Aberdeen

The reinforcing event is present as a stimulus in the situation in which a response is acquired and is absent from the extinction situation. When extinction conditions are made more similar to those of acquisition by introducing the reinforcer unrelated to behaviour ('free' reinforcement), the extinguished response is restored. Experiments with students, pigeons, and rats are described; and an analysis of the stimulus effects of reinforcement is used to account for 'disinhibition', 'spontaneous recovery', and speedy reconditioning after extinction.

I. Introduction

Reinforcing events are sources of stimulation, whatever else they may be, and therefore, in learning situations, they may be expected to gain some degree of stimulus control over the performance of the learned response. The experiments* that are described below demonstrate that renewed responding tends to occur when an extinction situation is made more similar to an original acquisition situation by adding the reinforcing event as a stimulus.

II. Experiments 1–4

1. *Restoration of extinguished operant in rats*

A conventional Skinner box was used. The floor space was 10 by 8 in. The manipulandum, a metal bar, was set at a height of 2 in. and the food cup was placed at floor level below the lever and slightly to the right. Responses were recorded by a Harvard Cumulative Recorder. (This recording instrument—used also in Exps. 2 and 3—was provided through a grant from the American Academy of Arts and Sciences.)

Six rats, deprived of food for 8 hr., were reinforced 20 times on each of 2 successive days of training by the delivery of one small food pellet following all bar-pressing responses. On the next 3 days each rat was put into the box for 30 min. extinction, lever-pressing now producing the usual noises but no pellets of food.

At the end of the last extinction period, when the frequency of responding was minimal, a number of stimuli were presented. They included tapping the box, flickering the lighting and sounding a buzzer. These stimuli had no detectable 'disinhibitory' effect. A few minutes later, a pellet of food was dropped into the food dish. All rats ate immediately and five out of six pressed the bar at least once during an observation period of 1 min. following eating. No rat had responded during the equivalent observation period before the delivery of food. Thus the 'free' issue of food had the effect of restoring the extinguished response.

Skinner (1938) mentions a similar case where the discharge of three pellets of food from the magazine by the experimenter resulted in some restoration of responding in an extinguished rat. He also reports a failure to obtain 'disinhibitory' restoration of operant responses when other stimuli were used.

2. *Restoration of extinguished operant in pigeons*

Six pigeons maintained on limited rations were trained to peck at a target set on the front panel of a Skinner box at a height of 8 in. The floor space of the box was 12 by 14 in. As reinforcement, one maple pea was delivered into a food dish at floor level below the pecking target for each correct response.

* Exps. 1, 2, 5(*a*) and 5(*b*) were part of an investigation supported by a research grant from the University of New Zealand.

After preliminary training with a lure, the pecking response was reinforced 50 times on each of 2 successive days. The response was then extinguished in three daily periods of 30 min. with the apparatus set as before, except that the food magazine was now empty.

At the end of the second extinction period, a number of tests of 'disinhibition' were made with a variety of stimuli. One of these, the rattling of maple peas, had a slight effect on the extinguished response in the case of two of the birds. The others had no effect. At the end of the third extinction period, one maple pea was dropped into the food pan at a prearranged time. All birds ate immediately, and five out of six pecked at the target within 10 sec. of eating. The bird that did not peck was given a second trial and responded then. No responses occurred in the equivalent periods before the free issue of food.

Thus rats and pigeons behave similarly when given 'free' reinforcement under similar conditions. In the case of pigeons, one stimulus (rattling of peas) may be slightly effective as a 'disinhibitor'. This is discussed later.

3. *Restoration of extinguished operant in students*

A slot machine was used. Each student was given forty counters and asked to aim at having the greatest possible number of counters in his possession at the end of the experiment. He was told that he could insert counters in the machine or not, just as he pleased, and that if he became bored he could read instead. Some old dry 'weeklies' were at hand.

Subjects of one group received total reinforcement. The first thirty-two responses of inserting a counter delivered two counters per response. Subjects of the other group were given irregular partial reinforcement. Of the first thirty-two responses, sixteen were reinforced by the delivery of four counters.

Extinction followed reinforcement without a break and continued until the subject had reached the criterion of making no responses during a period of 1½ min. A test of the effect of 'free' reinforcement then began. A control observation period of 30 sec. was followed by a 'free' delivery of two or four counters and a further period of 30 sec. observation. Of thirteen subjects, two responded during the control period while twelve responded in the test period immediately after the 'free' reinforcement. The partially reinforced subjects generally produced bursts of responses.

All of the subjects who responded acknowledged a connexion between their response and the 'free' issue of counters. Asked what this connexion was, they gave answers that fell into three groups. Four provided a rationale of some sort, for example, that the free issue was intended to give a hint or that it indicated that the machine had been refilled. Two explanations of this type were preceded by the qualifying phrase, 'I suppose I thought...'. Four subjects gave 'reasons' such as 'Well, I had some in my hand so I thought I would try again' or 'I thought I might as well, when I was given two more'. The remaining four fell into the 'don't know' category. On the whole, there was little to suggest that the behaviour was the outcome of reasoning in the normal sense of the word.

4. *Restoration of extinguished respondent in students*

Similar restorative effects can occur when the U.S. is presented alone after the extinction of conditioned respondent behaviour. Pavlov (1927) reported some observations by his collaborators using the response of salivation in dogs. An experiment carried out by Brebner (1957) at Aberdeen shows that the effect occurs with a different species and a different response.

In student subjects an eyelid response was conditioned to a tone of 1110 cyc./sec. Reinforcement was given on 60% of the acquisition trials using a puff of air to the cornea as U.S. After each subject had reached the criterion by making four C.R.'s on four successive test trials, the U.S. (puff of air) was entirely discontinued and extinction trials were given until four successive failures to respond to the tone had occurred. The U.S. was then presented alone on one occasion. Subsequently, when the C.S. was presented alone seven out of fourteen subjects responded. Thus, the restorative effect of 'free' reinforcement occurred as in operant conditioning although less markedly.

III. Interim Discussion

The experiments indicate that the restorative action of 'free' reinforcement is of some generality. In assessing the magnitude of the effect it should be remembered that 'free' reinforcement will strengthen whatever behaviour it happens to follow and, since this

behaviour is something other than the performance of the learned response, some counteraction of the restorative effect may have occurred.

The experiments might be regarded as demonstrating a type of 'disinhibition', but the theory that this term suggests may not be necessary and some additional hypothesis would still be required to account for the special effectiveness of 'free' reinforcement compared with other stimuli, as in Exps. 1 and 2.

As a first step towards explanation, the following empirical generalization may be useful: 'Maximal responding during extinction is obtained only when the conditions under which the response was reinforced are precisely reproduced' (Skinner, 1938). During extinction, the reinforcing event is necessarily absent. Considered as a stimulus it can be expected to have gained considerable control over the performance of the learned response, so that renewed responding will occur when it is restored to the extinction situation as a stimulus.

The principle provides a general framework for an account of the observations, but it leaves gaps to be filled in by a more detailed examination of the stimulus effects of reinforcers. For example, the stimulus of 'free' food may initiate a chain of activities with a corresponding chain of stimulation. The extinguished rat may eat the food, clean his whiskers, move around in the box and look at the bar. Giving free food may have quite remote stimulus effects, exteroceptive as well as interoceptive.

Some further experiments with pigeons may help to show what goes on. They deal with the restoration of responses originally set up by generalization rather than direct reinforcement. One procedure shows the effect in a very persistent form.

IV. Experiments 5 (*a*) and 5 (*b*)

5 (*a*) *Restoration of extinguished generalization in pigeons*

The subjects were thirteen birds from an experiment on sensory preconditioning that had given negative results (Reid, 1952). Birds from both control and experimental groups were used. No difference in the behaviour of the different groups was observed either in the original experiment or in the secondary experiment that is reported here.

The birds were trained in a Skinner box (floor space 12 by 12 in.) to respond to one signal and not to another. A buzzer and a light stimulus each of 1½ sec. duration were used. The response of pecking at a target 8 in. from floor level was reinforced by the delivery of one maple pea every time it occurred within 5 sec. of the onset of the positive stimulus. Trials were given at slightly irregular intervals averaging 1½ min., and only one response per trial was reinforced. (During training, reinforced responses commonly took place just after the stimulus had been presented rather than during the brief period of presentation.) Generalization to the negative stimulus was extinguished in two experimental periods. On a final test, the negative stimulus was presented five times and responses were recorded during periods of 5 sec. preceding and following its presentation. One bird responded before the stimulus on one occasion and three birds made single responses after. The effect of the five stimulus presentations can be indicated by a score of 2 (responses after, minus responses before).

On the next fifteen trials, the negative stimulus was presented for 1 sec. and simultaneously a maple pea was delivered into the food dish at floor level below the pecking target. The effect of the negative stimulus was then retested on five trials when it was presented alone as before. Again, one bird responded once before the stimulus, but eleven birds now responded after, making a total of thirty-four responses out of a possible fifty. As a result of the fifteen trials of 'free' reinforcement, the score indicating the effect of the negative stimulus increased from two to thirty-three.

During the 'free' reinforcement trials, two birds were observed to peck at the target *before* eating the food. They were excluded from the results quoted above, but their behaviour is worth considering. To the observer, the beginning phase of the chain of behaviour set off by the 'free' food is similar to the beginning of the pecking response. During training, birds normally peck at the target from a position directly in

front of it. The food that is delivered as reinforcement is picked up and swallowed and the bird moves away. Between stimuli there may be ritualistic pacing or turning in the cage or the bird may settle down quietly in one part. The onset of the stimulus generally occurs when the bird is not in the pecking position and its first effects are to produce orientation followed by movement into this position. When 'free' food is delivered into the dish the first effects are highly similar. The bird orients and moves forward in the cage. In the case of the two birds under discussion, these preliminaries may have been so strongly connected with the movement of pecking at the target that this movement became pre-potent even over eating. In other birds, occasional compromise responses were noticed, e.g. tapping lightly on the wall of the box between the target and the food dish.

During the test trials from which the main result of the experiment was obtained the negative stimulus was presented alone. Under these conditions, the stimulus sets off orientation and forward movement through its association with the delivery of 'free' food. Since no food is present during the test trials, the pecking response has no strong competitor. However, two birds that did not respond to the stimulus during the test showed vigorous incompatible behaviour of pecking at the empty food dish.

The interpretation of the main result of the experiment in terms of the chaining of responses is consistent with a failure to obtain the effect when the free food was delivered to different parts of the cage at random rather than to the food dish. Under those conditions, the negative stimulus on test trials produced a type of searching behaviour.

An alternative explanation of the effect of 'free' reinforcement should be mentioned, if only to be dismissed. During the fifteen trials of food delivery some birds were observed occasionally to peck at the target immediately after eating. It is quite probable that this pecking may at times have occurred when the maple pea was still moving down the pigeon's lengthy oesophagus. A need-reduction theorist might believe that direct reinforcement of the connexion between the negative stimulus and pecking could occur under these circumstances. However, this type of forward action of reward seems unlikely, since the behaviour of pecking after free food was never very frequent and died out in all cases before the test started. The number of pecks made by individuals on test trials was uncorrelated with the number of pecks that had occurred after 'free' food.

Pecking after eating food delivered as reinforcement usually occurs at the beginning of discrimination training as a generalization effect. When training is done with a lure and care is taken to avoid following this behaviour with food it dies out in about fifteen trials, showing that on balance it is not being reinforced by the eating that immediately precedes it. The revival of pecking after eating that occurs when 'free' food is given can be accounted for by an analysis of conditions of stimulation, but detailed consideration of the point would not be in place here.

5 (*b*) *Restoration of extinguished generalization—continuous test*

Fifteen trials of 'free' reinforcement were given in the last experiment, but the figure is arbitrary and the restorative effect could be expected to reach a maximum with some unknown number. If the interpretation that has been suggested is correct, too few trials would not produce a strong enough conditioned approach movement to the region of the target and food dish, while too many would result in strengthening the direct movement to the food dish so strongly that it would completely suppress the incompatible response of pecking at the target. The final experiment tries to circumvent this difficulty. From an anthropomorphic point of view, the procedure can be regarded as an attempt at muddling the birds as much as possible. (Animals may often have been confused by experimental procedures but rarely by intention.)

The apparatus and stimuli used were similar to those of Exp. 5 (*a*). Discrimination training was given to six birds until, on forty criterion trials, they responded to all twenty presentations of the positive stimulus within 5 sec., but did not respond more than once to twenty equivalent presentations of the negative (generalization) stimulus. (Three birds were tested with both stimuli arranged in an irregular sequence, and three were tested with the positive and negative stimuli presented in groups, but there are no obvious differences in their later behaviour.)

The response of pecking after 'free' food was eliminated in fifteen trials when food was given uncorrelated with any stimulus. The confusing procedure then began. In an irregular sequence the birds were presented with: (1) the positive stimulus (response reinforced), (2) the negative stimulus plus a free delivery of food, and (3) the negative stimulus alone (response unreinforced). Fifty presentations of each were given at the rate of 10 per day, the interval between each presentation varying irregularly around 1½ min.

Scores were calculated for each day's performance. They indicate the effect of the negative stimulus when presented alone. They are obtained as before by subtracting the number of responses made in the five-second periods preceding the stimulus from the number of responses made in the equivalent periods following the stimulus. Before the introduction of condition (2), i.e. before the test started, the average score for ten trials of the negative stimulus was 0·33. On the first ten trials of the new procedure it rose rapidly and was maintained throughout the whole of the period of testing with little decrement (Fig. 1). For the whole of the 50 trials the average score was 26·3 (27·6 − 1·3). Unfortunately, the experiment had to be broken off at the 50th test trial, but it is clear that the response to the negative stimulus produced by this type of 'free' reinforcement procedure was highly persistent. The response was not accompanied by a general breakdown of discrimination, pecking before the onset of the stimulus remaining minimal.

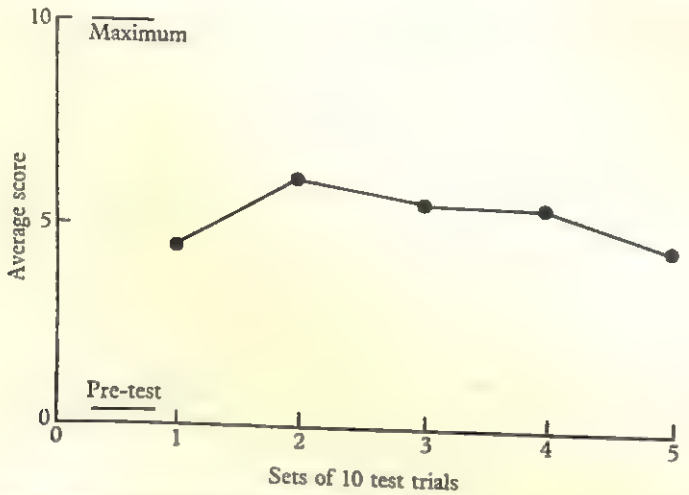


Fig. 1. Average scores for six birds on five sets of ten test trials. Horizontal bars indicate pre-test and maximum scores.

The confusing procedure ensures that reinforced pecking and eating are part of the bird's immediate past behaviour and that movement forward in the box in response to the negative stimulus is maintained in strength by the delivery of 'free' food in conjunction with this stimulus. Movement forward is an early part of a chain of behaviour that may carry through to pecking at the target when the stimulus of freely accessible food is not present to divert it. One bird was discarded for pecking at the target before eating 'free' food.

V. Discussion and conclusions

The results of these experiments can be understood through an analysis of the stimulus situations in which responses are made. The term 'set' might have been used in a descriptive way but there seems to be no need to isolate any group of stimuli as a class with special behavioural effects. The experiments are not put forward in support of any theory of reinforcement or extinction, but rather to suggest that when situations are closely examined, the need for theorizing may be reduced. Inhibition theories of extinction are widely held because the effects of extinction procedures do not appear to erase the effects of original learning. Demonstrations of 'disinhibition', speedy reconditioning after extinction, and 'spontaneous recovery' are taken to indicate that extinguished responses are somehow 'there all the time but inhibited'. It may not be necessary to think in these terms.

(1) '*Disinhibition*'

What is validly recorded for some purposes in a learning experiment as a single response may be part of a chain of behaviour. Chaining is most obvious in free operant behaviour, where, for example, a rat may press a lever, eat and press again repetitively. It is not surprising to find that Skinner (1938) discusses the matter in some detail. The regularity of operant responding may be partly attributed to the shortness of the cycles of behaviour that occur. In Pavlovian conditioning, chaining is less obvious because the stimuli are presented intermittently at long time intervals and bodily movement is restricted physically in the case of animals and by instructions in the case of human subjects. The occurrence of the C.R. may, however, be strongly influenced by the animal's state of vigilance and by its maintenance of posture (Guthrie, 1952; Razran, 1956). Although the experimenter records nothing between trials, this does not mean that the C.S. breaks in on a stimulus vacuum of any sort.

Pavlov found that in states of acute extinction the introduction of an extra stimulus would lead to renewed responding, provided that the intensity of the stimulus was not high enough to cause a strong 'defensive orientation reflex'. This is not surprising in view of the unfailing soporific effects of Pavlovian extinction. Any stimulus that revives vigilance and the posture that goes with it without emotional disturbance will restore some of the normal contextual stimuli of the C.S., so that when it is next presented a response can be expected. Since sleep was regarded by Pavlov (1927) as '. . . nothing but internal inhibition which is widely irradiated. . . ', any stimulus that is generally arousing could correctly be called a 'disinhibitor', but Pavlov seems to have envisaged a much more specific type of action. In the light of the present analysis it is understandable that the most effective 'disinhibiting' stimuli are those that make the dog salivate, especially the original U.S. Once started, salivation generally continues for a few minutes and may have much more lasting effects. For example, Pavlov reports an experiment on trace conditioning in which the U.S. is regularly presented to a dog every 30 min. If one presentation is missed the dog salivates at, or slightly later than, the 30th minute from the last presentation. The occurrence of the U.S. and U.R. at a given time can thus be registered and affect behaviour much later. When 'free' reinforcement is given in an extinction situation, the normal context in which the C.R. was established is most fully restored. The importance of the response as a source of stimulation may be indicated by the difficulty of extinguishing a C.R., the magnitude of which approximates to its U.R. (Razran, 1956).

In the Skinner box situation 'disinhibition' has proved to be difficult or impossible to demonstrate (Exps. 1 and 2, Skinner, 1938). Here the animal is free to move about, and it is less likely that any extra stimulus will restore the distinctive context of the acquisition trials. Merely alerting the animal may have little effect. It may in fact be actively preening or scratching or re-exploring its surroundings. 'Disinhibition' effects will normally be produced only when the stimulus starts off the early phases of the learned response. In the case of the free operant, giving free reinforcement re-establishes a chain. In the case of discrete responding (Exps. 5 (*a*) and 5 (*b*)) a stimulus that makes the animal move forward in the box may be effective for the same reason. In Exp. 2, the only stimulus that might be judged to have had a disinhibiting effect was the rattling of food, and this as a stimulus did in fact produce movement towards the food dish.

The disinhibition theory does not provide an alternative explanation of the relative difficulty of demonstrating extra stimulus restoration of responses in the operant conditioning situation, nor does it account for the special effects of presenting the reinforcer. Even the straightforward account of 'disinhibition' involves additional hypotheses to explain why extra stimuli affect inhibitory processes more than the excitatory processes they suppress, in Hull's (1952) terms, ${}_{S}I_{R}$ more than ${}_{S}H_{R}$. Konorski (1948), acknowledging the extra effectiveness of the U.S. as a 'disinhibitor', attributes this to 'inertia of the food centre', thus endowing a pseudo-neurological entity with pseudo-mechanical properties.

(2) *Speedy reconditioning after extinction*

Since reinforcement can never be withheld without altering the stimulus situation its reintroduction, even after complete extinction, can be expected to disclose some effect of the original training. The magnitude of the effect depends on the locus of extinction and this in turn is determined by the characteristics of the learned response as a chain.

When a response is made up of a number of units as a chain, the early units will be more remote from reinforcement than the later units and so can be expected to show less resistance to extinction. Further, in normal responding the occurrence of later units depends on the occurrence of earlier units. When the probability that an earlier unit will lead to a later unit is reduced below certainty, the later unit will in fact be performed less frequently than the earlier one and as a result will be less subject to extinction effects. Extinction procedures are unlikely then to extinguish every part of a response. As an example from outside the laboratory, if piano-playing ceases to be reinforcing, the break will probably take place at the stage of sitting down at the keyboard leaving the whole skill of playing intact. In general, the core of a skill is rarely exposed to extinction. The rat in a Skinner box may fail to respond when the early phases of his approach to the lever have been extinguished. Reconditioning will be fast because it does not involve setting up the whole response from scratch. An inhibition theory of extinction is not required.

(The stimuli that link a chain of responses need not be kinaesthetic. For a pigeon in a situation of the type used in these experiments, it seems likely that they would be predominantly visual. Even small movements in a Skinner box may be accompanied by large-scale visual changes. In a pilot experiment with four birds, the response of stepping on to a platform 2 in. from the floor was established. The platform was then removed and two birds developed the new response of standing where it had been and stretching their necks, presumably thereby reproducing the visual stimulation that was present at the time of reinforcement when standing on the platform.)

(3) *'Spontaneous recovery'*

It is regularly observed that the performance of an animal at the *beginning* of a second extinction period may be more similar to its performance at the *beginning* of the first extinction period than to its performance at the *end* of the first extinction period. This is called 'spontaneous recovery'. It is not surprising that it should occur since the situation at the end of one extinction period is different from the situation at the beginning of the next. This factor has been stressed by Skinner (1950) and Guthrie (1952). Skinner (1953) goes on to show that pigeons can discriminate their position on different schedules

of reinforcement in terms of elapsed time and number of responses made since last delivery of food. In experiments where trials are massed and responses are effortful, the difference in stimulus conditions between the beginning and the end of a period will be more marked and so more 'spontaneous recovery' will occur. During acquisition periods, reverse effects may occur. For example, a discrimination may be made with greater accuracy at the end of one period than at the beginning of the next. With pigeons these 'warm up' effects may be greatly reduced when the beginning of the experiment is made less distinctive by giving 'free' reinforcement before the stimuli are presented for the first time each day.

Finally, some of the concepts used in current theorizing are sources of confusion since they are remote from the details of behaviour. For example, the term 'habit strength' is used in an absolute way as if responses could be measured for strength like bars of steel. In particular, if the persistence of a response when reinforcement is withdrawn is taken as a measure of habit strength it should be recognized that the 'S' in ${}_{S}H_{R}$ during extinction is not the original 'S' of acquisition. The case where only one acquisition or extinction trial is given per session may be an exception, and it is interesting to note that under such conditions the concept of 'spontaneous recovery' has no meaning. Under more normal conditions of successive trials a test of the persistence of a habit is always a test of its generalization to stimulus conditions that are different from those of acquisition. The result can be regarded as a measure of the organism's discrimination. Thus, the finding that responses reinforced intermittently are more persistent than responses reinforced on every occurrence, gives no indication that response strength in any absolute sense is greater. The change in situation when all reinforcement is withheld is less marked in cases where reinforcement was sometimes withheld during the acquisition (Skinner, 1950).

The approach that has been outlined may deprive inhibition theory of some of its justification, but it is not strictly incompatible with the theory. In part, at least, inhibition theory may yet prove to be necessary. But it would be unwise to interpret extra stimulus recovery as a complex 'inhibition of inhibition' without first asking whether something simpler will do.

References

Brebner, J. M. T. (1957). An experimental investigation of the relationship between conditioning and introversion-extraversion in normal subjects. M.A. Thesis, University of Aberdeen.

Guthrie, E. R. (1952). *The Psychology of Learning* (rev. ed.). New York: Harper Bros.

Hull, C. L. (1952). *A Behavior System.* New Haven: Yale University Press.

Konorski, J. (1948). *Conditioned Reflexes and Neuron Organization.* Cambridge University Press.

Pavlov, I. P. (1927). *Conditioned Reflexes* (trans. G. V. Anrep). London: Oxford University Press.

Razran, R. (1956). Extinction re-examined and re-analysed: a new theory. *Psychol. Rev.* **63**, 39–52.

Reid, R. L. (1952). A test of sensory pre-conditioning in pigeons. *Quart. J. Exp. Psychol.* **4**, 49–56.

Skinner, B. F. (1938). *The Behavior of Organisms.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1950). Are theories of learning necessary? *Psychol. Rev.* **57**, 193–216.

Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. *Amer. Psychol.* **8**, 69–78.

(*Manuscript received* 22 *May* 1957)

VISUAL SEARCH FOR SUCCESSIVE DECISIONS

By N. H. MACKWORTH and J. F. MACKWORTH

Medical Research Council, Applied Psychology Research Unit, Cambridge

When a situation provides information from many different independent sources, characteristic difficulties arise, even when the decisions required are of the simplest possible kind and do not require estimates of probability. This paper considers the effect of the amount of material in the display which must be searched for the correct answer, and also how this display load interacts with the speed at which decisions are required, in a situation where successive decisions are posed at approximately equal temporal intervals.

It is shown that the percentage of errors is proportional both to the required speed and to the display load, and it is concluded that the time required to search through a display depends on the amount of material in the display, when this unwanted material closely resembles the required answer. In the discussion, the effects of many different sources of information in a moving display are considered in relation to two distinct variables: (1) the display load, and (2) the temporal irregularity of successive decisions.

I. Introduction

It is possible to study human responses to a complex environment in many different ways. For example: (1) In detailed controlled experiments the environment can be so simplified that as far as possible only one variable is being changed at a time in a limited situation; and with this approach a great deal of accurate information can be obtained about the effects of that single variable. (2) Alternatively, observational studies can be made of people working in their normal environment, in which case the findings are largely descriptive, since a given environment seldom repeats itself exactly. (3) Thirdly, however, simulated tasks provide an intermediate approach in which a fairly complex task is devised for use in laboratory experiments. This simulated task resembles the real situation in certain respects because both the real and the simulated forms of work contain common dimensions of difficulty, which can be altered in a controlled fashion. The interpretation of evidence from simulated tasks suffers from some lack of clarity in relation to the first approach mentioned above, and some lack of reality in relation to the second approach. The simulation experiment appears, however, to be an essential stage in bridging the gap between these two more conventional procedures, so that eventually the precise information obtained by very exact and controlled experiments may be applied with insight to the interpretation of observations from real situations.

A human problem of some importance occurs when action has to be taken on information coming at the same time from several different places in the environment. These sources may be quite independent or they may be related in some way. It may be easy or difficult to decide which source requires attention most urgently. Examples of multi-source displays from real life are found in many dial-watching tasks, a form of work which has been experimentally studied by Conrad (1951, 1954*a*, *b*, 1955*a*, *b*). Other instances, requiring still more difficult decisions, are the complex electronic displays that provide information to air traffic controllers. The positions of aircraft often have to be altered after these electronic maps have been carefully scrutinized. Success in such work may

at times be hard to attain, and, paradoxically enough, many of the difficulties arise from the presence of too much information in the display (Mackworth, 1949).

The following experiments are aimed at an understanding of some of the possible general reasons for these difficulties. It seems important to try to separate the main variables by devising new experimental situations. These laboratory tasks have to give exact and consistently reproducible results in circumstances which nevertheless avoid the dangers of constraining human actions to such an extent that the experimental situation involves only very simple forms of behaviour. The suggested method provides one way of analysing the factors arising when people are searching for a swift succession of decisions and are choosing these from a complex and crowded visual world. This information is quickly changing all the time for minutes on end; and it is often also relatively unambiguous, and does not necessarily entail probability judgements to any extent.

The average speed at which successive actions are demanded is not the only way in which work of this kind can become difficult. Two further variables have been isolated from such tasks—temporal irregularity and display load. Both these are characteristic of situations in which there are many sources of information. When a number of separate sources of information are generating stimuli independently at different rates, there are periods with many stimuli close together in time and these busy spells may be followed by long blank pauses (Cox & Smith, 1953). This temporal irregularity with its resulting tendency for the signals to overlap with each other along the time scale has been considered in a previous paper (Mackworth & Mackworth, 1956). The present work concentrates on two other variables—the average demanded speed of decision and the display load which is the amount of material simultaneously visible that must be searched for the correct decision.

II. METHOD

(a) *Nature of the order search task*

The experimental situation was designed to provide a mass of detail from which the subject had to pick out the one objective that was appropriate at any given time. As soon as this present objective had been identified by visual search, the action required on it was very simple and the next objective had to be sought (Mackworth, 1948).

The basic task was to bring a number of counters to the right places at the right time. Each man was tested by himself and sat in front of a visual display on which were shown: (1) a three-figure clock; (2) a large card marked off in half-inch squares; and (3) a set of counters placed on this grid. Each of these counters was labelled with its own reference letter, as in Fig. 1.

Each counter had to be moved down its own column on the grid. In this column there were four objectives, each marked with the time by which that particular square had to be reached. This expected time of arrival was given as a three-figure number in the same way as the readings on the clock.

At regular intervals the clock moved on one digit and ticked loudly whenever it did so. The subject had to give one order for each tick; he called out the index letter of the counter he wished to be moved so that the experimenter could move it on manually one square down the column. The subject was therefore searching for the forthcoming objectives in their correct sequence at a speed of work determined by the clock. He had one decision to take for each objective.

Fig. 1 shows an example of one of the displays. In this particular presentation there are twenty counters to be moved. As the clock ticks on from '000' the subject should name counter *I* twice for the first two ticks, as a result of his first decision to select the objective '002'. Then the counter *T* should be named three times after the second decision to pick objective '005'. Then he should name *M* repeatedly to achieve objective '015', and *Q* to reach '017'. Periodically some counters had to arrive early to ensure that subsequent counters arrived on time. Such early arrivals were not penalized in the present study. Thus the subject could not verify that he had the correct objective by counting the moves still required to

reach it; he had to consider every objective to make sure there was not another one even slightly earlier than the one found so far—unlike the more usual number search tasks. However, if the task was being performed correctly, the search was limited to the next objective in each column immediately beyond the counter.

The subject could obtain more time for his search by looking ahead. This involved searching for the next objective while calling out the moves already decided on to achieve the present objective.

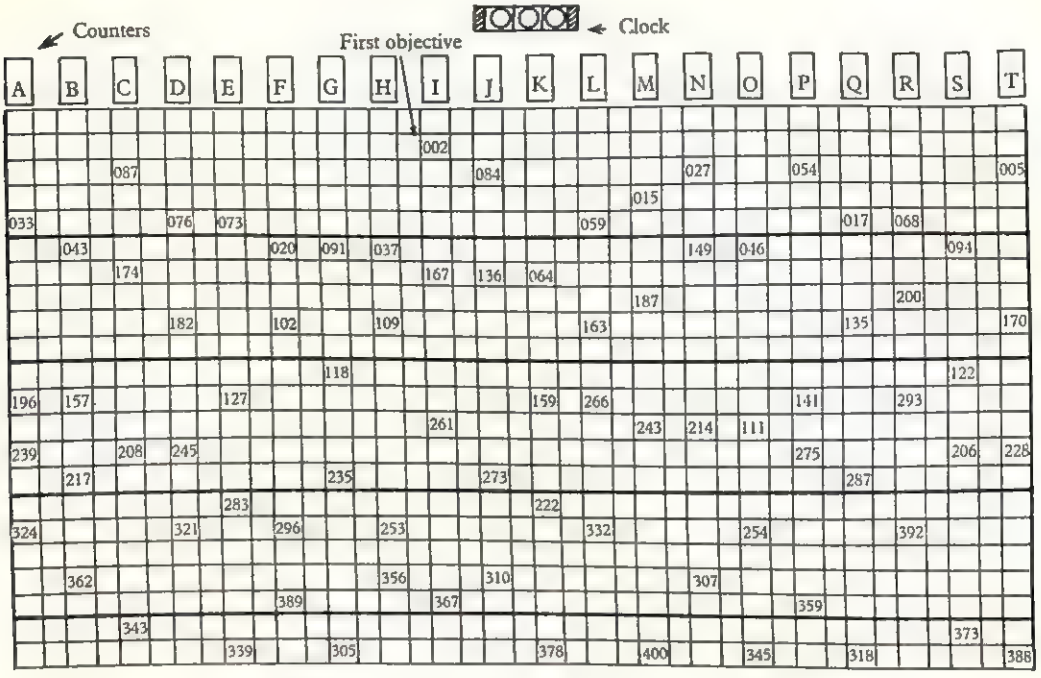


Fig. 1. Order search display, example of heavy load—20 columns.

(*b*) *Scoring method*

The method of scoring was to record, at each of the expected times of arrival, the difference between the expected and actual positions for that particular counter at that instant. This was expressed as the number of squares that the counter lay short of, or beyond, the objective. The experimenter dictated these numbers on to a tape recorder. The total error score was expressed in terms of errors per man per 100 moves, which represented 20 decisions. This was calculated by dividing the total errors per run by the number of moves per run (i.e. 100 moves for 5 columns to 400 moves for 20 columns).

By taking observations at each expected time of arrival, a semi-continuous record of achievement was kept during each run. It was specifically intended, for the sake of realism, that early errors should be allowed to prejudice later performance, i.e. that counters not in their correct position should further confuse subsequent action.

(*c*) *Experimental variables*

The variables studied were the *average demanded speed of decision* in terms of the average number of decisions per minute, and the *display load*, which for this experiment was defined as the number of columns through which visual search had to be made to find the present objective. Often, in practice, additional sources of information are forced into the environment without allowing the subject any extra time. It was desirable, however, to consider what happened when each of these two factors of required speed and display load was altered separately, as well as their effects when both were increased together.

For the first experiment, twenty conditions were considered since there were five demanded speeds of decision and four levels of display load. The five *speeds* were achieved by making the clock tick at different rates. The fastest speed entailed 12 decisions per minute, and the slowest only 2·4 decisions per minute.

In terms of moves of the counters, this meant a move every second or a move every 5 sec. (The search for a further objective had to be made within the time available for one move, unless the subject looked ahead and searched while he was calling out the present answer.)

The four *display loads* were obtained by increasing the number of columns presented. There were five, ten, fifteen or twenty columns. Each column always contained four objectives. (The columns were spaced out regularly across the card, so that the outermost columns always lay near the edge of the card, which measured 25 in. wide and 13 in. high, and subtended a visual angle of $\pm 35°$ horizontally and $\pm 13°$ vertically, since the eyes were about 20 in. from the top of the card.)

(*d*) *Experimental arrangements*

For the first experiment, the statistical design used five groups of four men. Each group did all the twenty conditions. The members of each group were allocated at randon to the four levels of load, but each man did each speed, in random order. Thus each man did five different conditions once.

The length of run for each display varied from 100 to 500 sec. for the five columns, and from 400 to 2000 sec. for the twenty-column display.

For the second experiment twelve conditions were tested, six levels of load, which ranged from five to fifty columns, at each of two speeds, viz. three and six decisions per minute. Each man did all the conditions in a Latin square arrangement. Six of these were in the morning and six in the afternoon.

Forty-four naval ratings, aged eighteen to twenty-nine, took part in this pair of experiments, twenty men for the first and twenty-four for the second. All were given an individual demonstration, followed by a practice spell of an hour an a half.

III. RESULTS

(*a*) *Results from Experiment* 1

Table 1 gives the average error score per 100 moves at each of the experimental conditions.

It can be seen from Table 1 that both the required speed and the display load have a considerable effect on the error scores. Fig. 2 shows the relation of display load to errors when the speeds were grouped into fast and slow, and Fig. 3 shows the relation of speed to errors when the display load values are grouped into high and low.

Table 1. *Average error scores per* 100 *moves in Experiment* 1

Speed (decisions required per minute)	Display load (number of columns) 5	10	15	20	Means
2·4	1	4	3	5	3
3·0	0	3	5	13	5
4·0	3	7	16	18	11
6·0	6	14	28	20	17
12·0	25	29	54	72	45
Means	7	12	21	25	—

Owing to the dependence of the variance upon the means in these results, it was necessary to apply a transformation to the readings. The simplest transformation which would remove this dependence was the use of the square roots of the errors. Table 2 shows the average square roots of the errors.

The analysis of variance on the group square root totals at the five speeds is given in Table 3.

This demonstrates the highly significant effect of the required *speed* on performance, and also that the differences between group totals at the same speed are not significant. This suggests that subject differences are not important when we take the average of four readings. The averages given in Table 2 are the averages of five results, so that an

analysis of variance is probably justified and this analysis of variance of the transformed data is therefore given in Table 4.

These results demonstrate the highly significant effect of *display load* upon performance.

It is important to check whether these differences obtained between the various display loads are due to the extra length of the task with the heavier loads. One way of studying

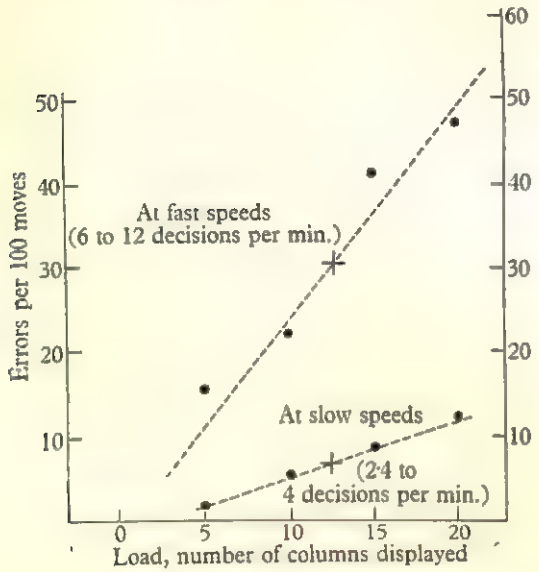


Fig. 2

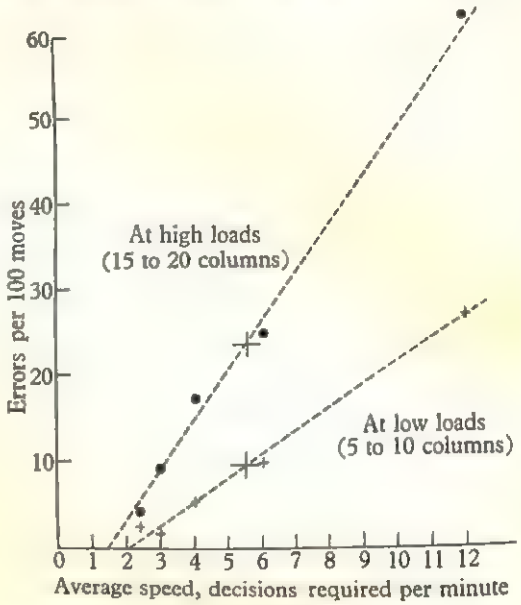


Fig. 3

Fig. 2. Effects of increasing the display load.

Fig. 3. Effects of increasing the required speed of decision.

Table 2. *Averages of square roots of errors in Experiment* 1

Speed (decisions required per minute)	Display load (number of columns) 5	10	15	20	Means
2·4	0·5	1·5	1·8	1·7	1·4
3·0	0·0	1·7	2·1	3·4	1·8
4·0	1·2	2·6	3·7	4·1	2·9
6·0	1·8	3·4	5·0	4·3	3·6
12·0	4·3	5·2	7·3	8·3	6·3
Means	1·6	2·9	4·0	4·4	3·2

Table 3. *Analysis of variance of group square root totals in Experiment* 1

Source	D.F.	S.S.	M.S.	*VR*	Significance
Speeds	4	90,323	22,581	41·8	<0·01
Groups	4	3,895	974	1·8	N.S.
Error	16	8,644	540	—	—
Total	24	102,862	—	—	—

Table 4. *Analysis of variance of Table* 2 *from Experiment* 1

Source	D.F.	S.S.	M.S.	*VR*	Significance
Speeds	4	59·7	14·9	41·4	<0·01
Loads	3	23·8	7·9	22·9	<0·01
Error	12	4·3	0·36	—	—
Total	19	87·8	—	—	—

this possibility is to consider only the errors in the first 100 moves for each of the various display loads. (This method of analysis *underestimates* any difference between high and low loads since, with the twenty-column load, the first 100 moves refer only to the first objective in each column, whereas with the five-column load the errors are cumulative through four objectives in each column.) Nevertheless, a regular increase in the average error score per hundred moves is again found, as can be seen in Table 5, which is averaged over the various speeds.

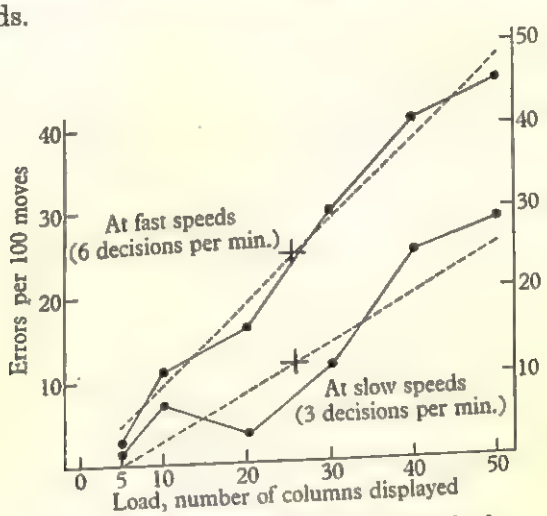


Fig. 4. Effects of greatly increasing the load.

Table 5. *Average error score during first* 100 *moves in Experiment* 1

Display load (number of columns)			
5	10	15	20
7	9	11	13

Table 6. *Average error scores per* 100 *moves in Experiment* 2

Speed (decisions required per minute)	Display load (number of columns)						Means
	5	10	20	30	40	50	
3	1	7	3	11	25	28	12·5
6	3	11	16	29	41	45	24·2
Means	2·0	9·0	9·5	20·0	33·	36·5	—

(*b*) *Results from Experiment* 2

Fig. 4 and Table 6 give the results of the second experiment in which the display load was greatly increased from five to fifty columns. The two lines on Fig. 4 show the errors per 100 moves for the two speeds plotted against the display load.

IV. DISCUSSION

(*a*) *The order search task*

In this laboratory situation the time-table numbers indicate the objectives for the search, and specify these in space as well as in time. The subject has to select these objectives in a sequence, and the clock indicates which of the objectives is the present objective at

any given moment. To bring each counter to the right place at the right time, the subject has to note the present clock reading and to calculate quickly the differences between this and the times given by the various possible objectives. As soon as he has found the objective that is the first one after the present clock reading, he calls for action on the counter aimed at that objective.

In other words, when he is making a decision, the subject starts with the differences in these time readings; and each of these time differences can be termed a temporal mismatch. MacKay (1956) has considered how a controlling system works from moment to moment in selecting what an effector system should do next, by choosing amongst the various possible lines of action. The effective stimulus is a mismatch. This is a difference between the desired and actual state of affairs, a mismatch which man is usually trying to minimize.

As in many real life situations, the main difficulty in selecting the correct action lies in the fact that all the possible objectives resemble each other. There is no difficulty in reading these three-figure numbers. The main problem is that the present objective can be identified only after the subject has rapidly considered a whole group of potential immediate objectives in relation to the present clock reading.

The nature of the task may change as mistakes accumulate, and the run proceeds. If the counters are being moved correctly, the task could be performed without looking at the clock, the subject merely having to find the next number in a sequence, but as soon as some objectives have been missed, then the correct move can only be found by considering and comparing the various temporal mismatches.

(*b*) *Relationship between speed and load*

(1) It has been statistically proved that both load and speed are important factors in the order search task. The display load has been defined here as the number of columns in the display through which visual search has to be made to find the present objective—and the speed is regarded as the average number of decisions required per minute. The result proved that error scores increase when the complexity of the display is increased in this way, even when the number of decisions demanded per minute is held constant.

(2) The results also suggested certain relationships which cannot be statistically proved on the present data. These indications are discussed below in paragraphs (3) to (7) inclusive, and need further experimental work.

(3) Performance in terms of errors per 100 moves would appear to be linearly related to the display load, provided the demanded speed of work is kept constant. Figs. 2 and 4 indicate that this linear relationship seems to hold good up to the display load of at least fifty columns, which was the highest load investigated.

(4) The results in Fig. 3 suggest that performance in terms of errors per 100 moves is also linearly related to the average demanded speed of work measured in decisions per minute, given that the display load is kept constant.

(5) Fig. 2 suggests that the effect of display load is greater at the higher speeds, and Fig. 3 that the effect of speed is greater at the higher loads. It would therefore seem that the effects of demanded speed and display load are not simply additive, i.e. the effect of increasing one factor intensifies the effect of increasing the other, to give diverging rather than parallel lines in Figs. 2 and 3.

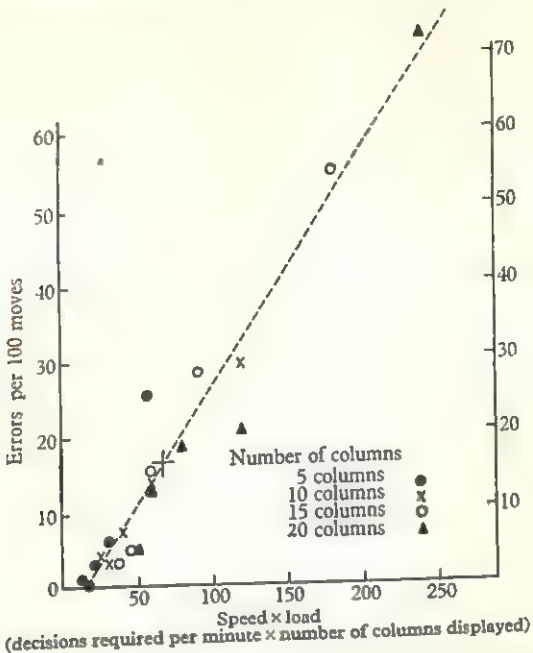


Fig. 5. Relation between errors and product of speed and load

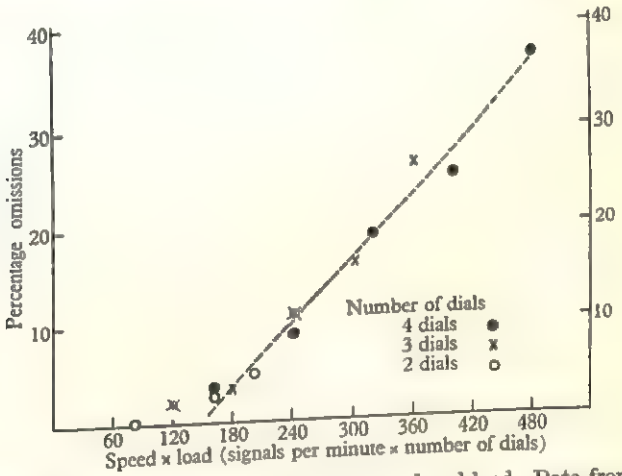


Fig. 6. Relation of percentage omissions to product of speed and load. Data from Conrad (1951).

(6) Inspection of the results shown in Table 1 suggests that in the circumstances of the experiment it is possible to obtain the same error scores when the speed is doubled and the load is halved. When the product of speed and load (as measured in the way stated) is constant, the error score is also fairly constant. (Approximately the same errors are obtained, for example, at a speed of 3 decisions per minute with a load of 20 columns as are found with a speed of 6 decisions per minute and a load of 10 columns. Similarly, an almost identical score is found with a speed of 4 decisions per minute and a load of 15 columns.) The relationship remains true even when, as in Table 2, the individual results are transformed into square roots to give a more normal distribution.

(7) These results suggest that the errors per 100 moves might be linearly related to the product of speed and load, and this relationship has been plotted in Fig. 5.

(8) Similar relationships can be found in the literature, e.g. Conrad (1951) in his work on multiple dials obtained results which have been re-calculated and shown in Fig. 6; this plots the product of demanded speed and display load (in terms of decisions per minute and number of dials) against percentage omissions. A straight line can be fitted to the data when the percentage omissions are greater than about 2%.

(c) *Search time and load in static displays*

It is thought that the time to find an objective depends directly on the number of objectives that must be inspected, because the subject takes a certain time to examine each objective and reject it. Eriksen (1955) has shown that when searching for ten objects, namely triangles, amongst other different shapes, such as squares and diamonds, the time to find the triangles was directly proportional to the number of other irrelevant shapes in the display, which contained a total of 20, 40, 60 or 80 shapes.

Similarly, Green & Anderson (1956) have noted that the time required to find a given numeral in a display is directly proportional to the number of other numerals in the display. It is therefore reasonable to suppose that a similar effect is occurring in the order search task, and that errors will occur when the pacing is too fast for the subject to find his present objectives. The main point is that failure is just as likely to arise from greater display complexity with more details to discard, as from demanding more decisions per minute.

(d) *Relation between the present objective and the rest of the display*

Some situations need hardly any search at all, as when a bright light is presented against a neutral background—e.g. a searchlight on a dark night. At the other extreme, no amount of search will reveal the wanted signal when this is identical with other unwanted parts of the display, e.g. when a weak target signal on a radar tube is quite indistinguishable from unwanted random flashes. Between these two extremes there are many instances in which visual search is rewarding but difficult. Such difficulties arise when the unwanted items in the display are more numerous than the present objective, and also resemble it in certain respects.

The unwanted items are not necessarily totally irrelevant in the sense that they are never going to be required. For example, in a serial task such as the order search situation, all the information shown on the display is relevant at some stage during the problem. At any given moment in the task, however, most of the information shown has to be discarded by the subject, not because it is entirely random information, but rather because it concerns past objectives or objectives which are far into the future. This has been shown to be harmful to performance. It is therefore important to remove from displays all the information that is not needed for present action. The less the predictive behaviour entailed by the particular series of decisions, the greater is the need for a display showing only the present objectives (Mackworth, 1956). This implies: (i) the removal of past objectives, (ii) reducing the number of future objectives, as well as (iii) the traditional elimination of 'noise' or random information. Electronic computers can help by storing information as it comes in, and by displaying only present objectives, these being shown just before the operator has to consider them.

Even if, however, it is not possible to have one of these 'now' displays, much can be done by some special labelling for the present objectives. Surprisingly little is known about the factors that lead to confusion between visual objects; it is however clear that the difficulty of search is related to the similarity between the present objective and the rest of the display material. For example, Green & Anderson (1956) have demonstrated that differential colour labelling of numerals effectively reduces visual search to those numerals in the given colour.

(e) Spatial search and temporal irregularity in multi-source displays

It is believed that there are two main sources of difficulty characteristic of a multi-source display (apart from the average speed required); these are (1) temporal irregularity, and (2) display load. Temporal irregularity has been reduced to a minimum in this order search task, while the effect of varying the display load at different required speeds has been studied in some detail.

On the other hand, there are undoubtedly a great many real-life situations in which signals are displayed independently from many sources so that both factors are present. The laboratory situation mentioned earlier, devised by Conrad (1951), presents a task involving both temporal irregularity and varying display load (Conrad, 1955*b*). With temporal irregularity, a marked increase in the demanded speed of work tends to reduce the disadvantage of increasing the number of channels, which is at a maximum at an intermediate speed. This can be seen in Fig. 4 of the paper mentioned above on signal overlap (Mackworth & Mackworth, 1956). Raising the demanded speed of work has exactly the opposite effect in the present laboratory task which involves searching the display; faster pacing increases the disadvantage of more channels. Thus in a complex task, the exact effects of demanded speed on multi-source displays may depend on the relative importance of temporal irregularity and display load.

This discussion has considered the relation between speed and load in terms of an error criterion averaged over a given number of moves or decisions. Quite a different picture is obtained when the score is measured in terms of the average number of responses *per unit of time*. Conrad (1954*a*) has pointed out that increasing the display load reduces the total output of correct responses per minute—whereas increasing the speed by demanding more decisions per minute may have the opposite effect of increasing the output of correct responses per minute.

Inevitably, therefore, when the measure of performance is the output of correct responses *per minute*, it is impossible to compensate for an increase in display load by decreasing the required speed, since both these changes tend to reduce the output per minute.

The importance of avoiding unduly heavy display loads is greatest in situations where it is essential that every decision must be made accurately at the right moment. Such a situation is more likely to be found in, for example, air traffic control than in most existing industrial situations where the yardstick is productivity per unit of time.

V. Conclusions

This experiment is a contribution to an analysis of the effects of increasing the number of sources from which the subject is expecting visual information about which he must decide. Consideration of the difficulties produced by increasing the number of sources has identified two main causes of failure, namely temporal irregularity and display load—apart from any other effects due to an increase in the average number of decisions demanded in unit time.

In a situation in which *temporal irregularity* is the limiting factor for performance, the main obstacles to success are believed to be the overlapping of signals for decisions during very active periods, followed by enforced pauses in which no action is required. Search problems are at a minimum since all the signals are easily seen and must be responded to as soon as possible after they appear. There is no need here for the concept of display load, because performance decrement can be accounted for by the moment-to-moment variations in the demanded speed of work. These variations are superimposed on the general average demanded speed.

On the other hand, *display load* may be the limiting factor when the display is a complex one, and there is difficulty in deciding which is the present objective. This is intensified when there is an increase in the difficulty of distinguishing between the wanted and unwanted information. The concept of display load is essential in describing visual search; in the task considered in this paper, *display load* has been defined as the number of columns through which visual search must be made to find the present objective.

It has been shown that errors per 100 moves in the order search task depend directly upon the display load as well as upon the speed. Moreover, it seems that the effects of display load tend to increase as the speed increases. This is in opposition to the findings with temporal irregularity, where increasing the average speed beyond a certain point tends to diminish the effects of temporal irregularity introduced by extra sources.

When accuracy is all-important to the visual search task and when percentage error is the method of scoring, it is possible to demonstrate some convertibility in the speed and load factors, in the sense that halving the speed may compensate for doubling the load, i.e. halving the number of decisions required per minute may compensate for doubling the number of columns on the display. It should be pointed out that since the speed in terms of the number of decisions demanded per minute is a product of the number of sources and the speed at which each source is requiring decisions, then whenever the number of sources is increased, the rate at which decisions come in *from each source* must be reduced very considerably to hold constant the product of average speed and load.

These researches owe a great deal to the encouragement and comments received from Sir Frederic Bartlett, F.R.S., as well as to the advice obtained from our colleagues during many discussions. The statistical design of the experiments was made by Mr E. G. Chambers and the analysis was by Mr Mervyn Stone.

VII. References

Conrad, R. (1951). Speed and load stress in a sensori-motor skill. *Brit. J. Indust. Med.* **8**, 1–7.

Conrad, R. (1954*a*). Speed stress. Chapter 13, in *Human Factors in Equipment Design*. London: H. K. Lewis and Co. Ltd.

Conrad, R. (1954*b*). Missed signals in a sensori-motor skill. *J. Exp. Psychol.* **48**, 1–9.

Conrad, R. (1955*a*). Timing. *Occup. Psychol.* **29**, 173–81.

Conrad, R. (1955*b*). Some effects on performance of changes in perceptual load. *J. Exp. Psychol.* **49**, 313–22.

Cox, D. R. & Smith, W. L. (1953). The superposition of several strictly periodic sequences of events. *Biometrika*, **40**, 1–11.

Eriksen, C. W. (1955). Partitioning and saturation of visual displays and efficiency of visual search. *J. Appl. Psychol.* **39**, 73–7.

Green, B. F. & Anderson, L. K. (1956). Color coding in a visual search task. *J. Exp. Psychol.* **51**, 19–24.

MacKay, D. M. (1956). Towards an information-flow model of human behaviour. *Brit. J. Psychol.* **47**, 30–43.

Mackworth, N. H. (1948). Researches on the measurement of human performance. Research Fellowship Thesis, St John's College, Cambridge.

Mackworth, N. H. (1949). Human problems of work design. *Nature, Lond.*, **164**, 982–3.

Mackworth, N. H. (1956). Work design and training for future industrial skills. (Sir Alfred Herbert 1955 Lecture.) *J. Instn Prod. Engrs*, **35**, 214–40.

Mackworth, J. F. & Mackworth, N. H. (1956). The overlapping of signals for decisions. *Amer. J. Psychol.* **69**, 26–47.

(*Manuscript received* 13 *May* 1957)

THE ROLE OF ATTITUDES IN COMPREHENSION

By S. FRIEDMANN

Child Guidance Clinic, City of Oxford

By means of an attitude test and expressed opinions, subjects were divided into pro-, anti- and neutral-groups for each of twenty-four political statements. No significant differences in comprehension were found between the three groups, except in two cases where the groups who scored higher for comprehension also had superior intelligence and knowledge scores. A significant result was found for only one of the twenty-four statements, namely, a negative correlation between liking and comprehension, and a comparable superiority in comprehension of the opposing group: this statement was found to be ambiguous in its expression of attitude.

The findings of a control experiment, using comments instead of a comprehension test, supported the main conclusion that for unambiguous statements there was no relationship between the efficiency of comprehension and direction of attitude.

I. Introduction

In spite of the growing interest in the relation between affective and cognitive processes, there have been very few experimental studies on the influence of attitudes on comprehension. It is often held that in order to understand an idea one must be in sympathy with it (Baumgarten-Tramer, 1941; Hippius, 1936; Hodges, 1944; Marum, 1936), or at any rate, that favouring an idea leads to better comprehension than does antagonism to it. On the other hand, this view has been strongly attacked on the grounds that understanding depends on knowledge and skill, and that favourable attitudes are quite irrelevant (Ryle, 1949). It might equally well be the case that neutral attitudes, or a critical approach, lead to better appreciation. The issue can only be decided experimentally.

Experimental studies of the relation between affective factors and cognitive processes have been mainly in the fields of learning, recognition, and remembering. There are several investigations of social attitudes—in the sense in which Sherif & Cantril (1945, 1946), used the term—which purport to show that favourable attitudes lead to better learning and remembering than unfavourable attitudes. Edwards (1941) tested the recognition of items in a New Deal speech of subjects with pro- and anti-New Deal attitudes. The favouring groups scored consistently better than antagonistic groups, although the differences were not statistically significant. Seeleman (1940) used two groups of subjects with pro- and anti-Negro attitudes in the recognition of Negro and White pictures. The anti-Negro group recognized fewer Negro pictures, and fewer individual differences between Negroes, and remembered more unfavourable than favourable phrases connected with Negro pictures than did the pro-Negro group. Levine & Murphy (1943) found Communists superior in learning and remembering a Communist passage to anti-Communists, while the latter scored better on an anti-Communist passage. Most of these investigators interpreted their findings in terms of Sherif & Cantril's concept of reference frames. It was contended that material which is congruent with the individual's frame of reference is more easily assimilated than material which is not.

Since learning, recognizing, and remembering are involved in understanding, these experiments might be deemed to build up a *prima facie* case for the contention that favourable attitudes lead to better comprehension. There is little direct evidence,

however. In a study of the intelligibility of spoken messages, Mason & Garrison (1951) obtained significantly higher scores for liked than for disliked sentences. Inspection of the material suggested to them that liking was more a matter of greater familiarity than of favouring. Cooper & Jahoda (1947) held that 'it is difficult in general for a communication to reach people who are not already in favour of the views it presents'. Subjects with anti-minority views misunderstood a series of cartoons ridiculing anti-minority feelings, and misconstrued the point so as to fit it in with their own *convictions*. This may well be a special case, however, since the subjects had to see themselves as ridiculous *if they understood* the cartoons. Unfavourable attitudes can be aroused quite strongly without necessarily threatening the subject's self-esteem. The results, therefore, do not necessarily support Cooper and Jahoda's more general hypothesis. Doob (1947) criticized experiments on social attitudes in learning, recognizing and remembering for not controlling for previous knowledge of the subject-matter. The findings might reflect no more than the better knowledge of the favouring groups. Alternatively, the results could be explained by assuming greater interest—hence better incentives—in the favouring groups. Thus, there is but slight evidence on which to base a theory of the relation between attitudes and comprehension. Most investigators demonstrate that the reproductions and interpretations of material reveal the subjects' attitudes, but it is by no means clear what functions these attitudes have, or how they are related to the efficiency of comprehension. The present investigation sought to determine whether there is any relation between the efficiency with which subjects understand statements, and their attitudes to the opinions expressed in them.

II. EXPERIMENTAL MATERIAL

Short political statements from publications of the major British parties were considered the most suitable material, since they are designed to arouse the kind of attitudes here to be investigated. In order to see if any possible effect of attitude on comprehension is connected with the level of difficulty of the material, the sentences were to be at three levels—easy, average, and very difficult. Over a thousand typical statements were collected from Conservative, Liberal, Labour and Communist party pamphlets. The experimenter and two other observers selected eighty of these as falling most consistently into five grades of difficulty. Sixty-two psychology students then ranked the eighty statements for difficulty. A further sixty-five subjects with a wide range of political opinions grouped the statements according to the four political parties, to ensure that, although taken out of their context, the sentences still expressed the views of the parties from which they issued. The two statements which ranked consistently in the first, third, and fifth grades of difficulty were selected from each of the four attitude groups. These twenty-four statements, printed on separate cards, were used in the investigation (see Table 1).

Table 1. *Statements for Comprehension Test (Material used in* 1950)

Easy statements

No. 36F. (Conservative.) Thousands of men and women would work harder today if they could look forward to buying a house with their savings.

No. 19L. (Conservative.) Working men are not frightfully anxious about having their teeth yanked out for nothing, or their children kept at school until they are of marriageable age.

No. 17D. (Liberal.) The Election Landslide of 1945 came as a surprise to many voters; both to those who voted for the Labour Party and those who voted against it.

No. 71N. (Ambiguous Liberal.) The basic fact about the Liberal Party is that it is opposed to both the present major parties.

No. 104B. (Labour.) In April 1928 the Tories reduced the State grant towards milk for expectant mothers.

No. 104H. (Labour.) On the third anniversary of Labour's General Election triumph, 5 July 1948, the full social security scheme was in force, the Family Allowances, the National Insurance, Old Age Pensions (increased to 26*s*.), and National Assistance.

Table 1 (*cont.*)

No. 39C. (Communist.) During the war, in every country that was fighting fascism the Communist Parties played a leading part.

No. 56H. (Communist.) The Labour Government carry out a purge against Communism in the Civil Service, but allow Mosley to speak in schools controlled by Labour Councils.

Statements of average difficulty

No. 52C. (Conservative.) Private property is a natural right of the family and generally of the subordinate group.

No. 68A. (Conservative.) As the Conservative sees it, the difficulty of private enterprise lies not in Socialist doctrine but in the size of unit created by modern technical development.

No. 71L. (Liberal.) I do not maintain that an effective union of Europe depends on every government being liberal, but they must not be illiberal.

No. 61D. (Liberal.) According to the orthodox Liberal view, the cause of wars was not Capitalism, but authoritarianism and in particular the authoritarianism of aristocracies and princes.

No. 73G. (Labour.) Under Socialism we shall be able to do with many fewer bureaucrats because the main industries will be organized directly as public services for the common benefit.

No. 63C. (Labour.) There are big industries not yet ripe for public ownership which must, nevertheless, be required by constructive supervision to further the nation's needs.

No. 100A. (Communist.) If the electoral programme of the Labour Party in Britain had been carried out in the interests of the working class it would have resulted in real inroads being made into the power of the capitalist class.

No. 40P. (Communist.) The bourgeoisie has subjected the country to the rule of the towns.

Difficult statements

No. 54F. (Conservative.) Conservatives think Liberals failed because they were afraid to anchor their theory of government upon the moral precepts of the natural law and sought instead something more objective which would be more universally recognized.

No. 29A. (Conservative.) The truth of the matter is that European recovery depends not only on preferential arrangements but on a revival of the Sterling Area and the creation of a dollar pool.

No. B1. (Liberal.) Under a well-constituted, or even under a well-administered though ill-constituted government, men's moral sensibility is commonly stronger, and their moral biases more conformable to the dictates of utility.

No. 67S. (Liberal ambiguous.) Inspired with the sacred writings of Adam Smith among economists and of Jeremy Bentham among political theorists, the Liberals of the nineteenth century proclaimed a jehad against all forms of government interference in the realm of commerce.

No. 6F. (Labour.) Only the acceptance of the major changes carried out through the recent Parliament would enable Conservatives to put over effective propaganda to many of the groups which are drawn from those sections of the electorate which in the past supported the Conservative Party as a 'safe' party.

No. 65V. (Labour.) Exactly in proportion as Conservatives believe in federation they distrust the passionate young nationalisms they have largely helped to create.

No. 38B. (Communist.) For many a decade past the history of industry and commerce is but the history of the revolt of modern productive forces against modern conditions of production, against the property relations that are the condition for the existence of the bourgeoisie and its rule.

No. 33E. (Communist.) The proletarians cannot become masters of the productive forces of society except by abolishing their own previous mode of appropriation.

III. Comprehension test

'An individual comprehends, or understands a situation', was considered to mean that the individual is able to vary his responses to the situation and can adapt them to variations in the situation in accordance with his purposes. While this may not cover all the senses in which the term 'understands' is used, it would appear to cover all those with which the present investigation is concerned. The behaviour referred to above is exhibited in general by demonstrating verbally or otherwise how a situation hangs together, how it works, what follows from it, what conditions precede it, and how similar new situations will work. The specific tasks involved in understanding statements have been given in the studies of Richards (1930) and Davis (1942, 1944). In order to make these tasks explicit, a comprehension test was constructed in the following form. (i) Explain in detail, wherever possible with concrete examples, in as SIMPLE terms as possible, what this statement means. Also explain any vague or ambiguous words the

statement may contain. (ii) Give reasons why you believe the statement to be true or untrue. (If you have insufficient information to judge whether the statement is true or false, state what kind of facts would make it true or false. In statements that state what somebody believes, give reasons BOTH for thinking whether they believe it, and whether the belief is true or false.) (iii) Is there anything the writer of the statement left unsaid, but which is implied in the statement? If so, what is it? (iv) What effect did the writer of the statement wish to produce on his reader? These questions were duplicated on foolscap sheets, leaving half a page between questions so that answers could be given in essay form (see Vernon, 1950). An observer who answered these questions satisfactorily for a given statement was held to understand it. The satisfactoriness of the responses was judged by three independent examiners who subsequently compared their criteria and pooled the results for the final scores. At the end of the form, space was left for introspective remarks with instructions to the subjects to comment fully on their reactions to the statement, and to give an account of what passed through their minds during the period in which they were trying to understand the statement.

IV. ATTITUDE TEST

'*A* has an attitude to *Y*', was held to mean that *A* tends to act and feel with regard to *Y* in an acquired but characteristic manner. Attitudes were distinguished from interests. '*A* is interested in *Y*', was held to mean that *A* tends to be motivated to attend to *Y*, deal with it, or respond to it. It was considered that attitudes describe the direction, consistency and characteristic manner of the organism's responses. This applies particularly to social attitudes, i.e. attitudes towards values and ideas held in a society. According to Bartlett (1932), such attitudes may be aroused immediately on first general impression, before the material is completely understood, and influence perception. To ascertain whether there is such an influence on understanding statements, subjects were asked the following question immediately after reading a statement. 'Did you like or dislike the attitude expressed or implied in the statement? Indicate your liking or dislike by putting a cross against one of the following: (i) Liked very much; (ii) liked; (iii) moderately liked; (iv) neutral; (v) moderately disliked; (vi) disliked; (vii) disliked very much.' (The ratings were checked to ensure that liking or dislike of attitude, and not of ease or difficulty, had been recorded.)

The observer's general political orientation was ascertained by an attitude scale, and by opinions expressed during the testing and in the scripts. The attitude scale was not envisaged as a linear radicalism-conservatism continuum, but consisted of twelve easy statements expressing Conservative, Liberal, Labour and Communist opinions, respectively. The subjects ranked these in order of preference.

V. POLITICAL KNOWLEDGE TEST

A political knowledge test was constructed, consisting of factual questions about British politics. Fifty-five questions on the names of ministers, parties in power, political literature, and figures of production and consumption were arranged in the form of a four-choice test, and given to eighty Arts, Science and Law students. An item analysis was performed, and the items were scaled for difficulty. Thirty-two questions were selected from those showing a higher correlation with the test and presenting an ascending order of difficulty. The items were so arranged that not too many difficult questions followed each other for the final form of the test.

In addition the N.I.I.P. group test 33 (verbal intelligence) was used.

Taking twice the standard error of the correlation (0·29) as the criterion of significance, the correlation of 0·48 between the N.I.I.P. test and the political knowledge test was considered satisfactory. Correlations between comprehension and the N.I.I.P. test were rather lower than had been hoped (0·30 to 0·50). In five cases, they were somewhat below the criterion. The statement which had been the first to be marked for comprehension showed low correlations both with intelligence and knowledge. Apart from this, the correlations between comprehension and political knowledge were well above the criterion of significance (0·34 to 0·63).

VI. PROCEDURE

Forty-nine subjects took part in the experiment. With the exception of six subjects, they were Arts and Science graduates and undergraduates. Each subject came for five sessions of approximately 1½ hr. Six statements were dealt with in each of the first four sessions, given in the order: easy, average, difficult, average, difficult, easy. In every group of six, there was at least one statement from each of the four attitude groups. After reading the sentence, the subject recorded his like or dislike of the attitude

stated or implied. Then followed the four comprehension questions, and finally a request for introspective remarks.

No time limits for reading or answering were imposed on the first nineteen subjects. It was found more convenient to set time limits for subsequent subjects. These were arrived at by slightly extending the average times of the first nineteen subjects, so that every one would have ample time to finish. Since there were no significant differences in comprehension between the two groups, they were treated together in the analysis.

In the final session, the intelligence, knowledge, and attitude tests were given, in that order. No time limits were imposed for the last two tests. The mean score on the N.I.I.P. test was 156·3 with a standard deviation of 19·3. The mean score on the political knowledge tests was 16·3 with a standard deviation of 5·2.

Most of the subjects were tested individually, some in twos, and there was one group of five subjects.

In order not to prejudice the results, the subjects were told initially that the experiments were made to validate tests. At the end of the sessions, their actual purpose was explained.

VII. Supplementary experiment

In case results were affected by the comprehension test which directed attention to various aspects of the statements, a control group of twenty subjects was asked to comment on nine of the statements instead of answering comprehension questions. Eight were medical and social science students, and twelve students of a University Extension class in Social Psychology. The mean score on the N.I.I.P. test was 147·7, and on the knowledge test 17·4. They thus came within the same intelligence and political knowledge category as the main group.

VIII. Results

Four hypotheses may be held about the relation between attitudes and understanding: Comprehension is better when (*a*) subjects' attitudes are in agreement with the views expressed in the statements; or (*b*) subjects are opposed to these views; or (*c*) subjects are neutral in their attitudes to these views; or (*d*) the efficiency of comprehension does not depend on the agreement, disagreement or neutrality of the subjects' attitudes to those expressed in statements of opinion. These four hypotheses were to be tested with regard, first, to the immediately aroused like or dislike of the content, and secondly, to the subjects' general orientation. In the detailed analysis of each statement necessitated by marking for comprehension, two Liberal statements 71N and 67S, were found to be ambiguous in attitude, i.e. they could be regarded as anti-Liberal also. The results for these two statements will be dealt with separately.

(i) *Immediate attitude and comprehension*

The seven point scales of liking to dislike were correlated with the comprehension scores for every statement. Intelligence and political knowledge were held constant by partial correlation. For this, all test scores were converted into standard scores and divided into seven groups, from best to worst, by 0·7 σ units ranging equally about zero from +2·5 to −2·5. (This was found to be the best approximation to a normal distribution. For statistical details, see Friedmann, 1952.)

None of the partial correlations for the twenty-two unequivocal statements were significant by the criterion—twice the standard error of the correlation. The distribution of the correlations was of a random order tending to zero. Since no positive or negative linear relationship, and no tendency in either direction, was found between immediately expressed liking and comprehension when intelligence and knowledge were held constant, the first two hypotheses must be regarded as disconfirmed for our material.

The third possible relationship between immediate liking and comprehension is a non-linear one. The correlations between liking and comprehension were, consequently, tested for linearity. With $k-2$ degrees of freedom (i.e. 5 for these cases) χ^2 is significant at the 0·05 level if it has a value above 11·07. (Fisher, 1950, table III). For twenty of the twenty-two correlations the deviations were not significant by this criterion, while two were significant at the 0·05 level. According to Wilkinson (1951), and Sakoda, Cohen & Beall (1954), up to three results, significant at the 0·05 level, can occur by chance in a sample of twenty-two. The two significant deviations can, therefore, best be regarded as chance results. This is supported by the internal evidence: a comparison between the scattergrams for comprehension and attitude, and those of knowledge and intelligence and attitude revealed no differences. It is, therefore, likely that for these two statements the most intelligent and knowledgeable subjects happened to be distributed in a particular way with regard to attitude. The scattergrams did not, in any case, show any tendency to be inversely U-shaped, as would be expected if neutral attitudes were connected with better, and like and dislike with worse, comprehension.

(ii) *General political attitude and comprehension*

Subjects were divided into three groups (i.e. favouring, opposing, and neutral) for every statement, on the basis of the attitude test and the opinions expressed in the scripts. The differences between the scores of the three groups were tested by Fisher's t-formula to determine whether there was any consistent or significant superiority. Any one of the first three hypotheses would be supported only if one of the three attitude groups showed significant superiorities in comprehension which could not be explained by similar advantages in intelligence or knowledge scores alone.

Among the sixty-six comparisons between comprehension scores, there were two significant results at the 0·05 level. Statistical considerations make it likely that these two results should be regarded as due to chance. The internal evidence supports this, since the groups that scored significantly more for comprehension also had superior intelligence and knowledge scores. Similar arguments also apply to the four significant results obtained among the 132 comparisons between intelligence and knowledge scores. The obtained differences for comprehension showed no consistent tendency in the direction of any one of the three attitude groups.

(iii) *Results for the ambiguous statements*

The sole criterion for dividing the subjects into favouring and opposing groups for the two ambiguous statements was the subjects' agreement or disagreement with the attitude they had taken the statement to express.

For the difficult statement 67 S, there was no difference between the favouring, opposing and neutral groups for comprehension, intelligence or knowledge, and the partial correlation between liking and comprehension was not significant.

There was a significant partial correlation ($-0{\cdot}3535$) for the easy statement 71 N. The correlation between attitude and comprehension departed significantly from linearity (χ^2 from $\eta\ CA = 12{\cdot}65$; χ^2 from $\eta\ AC = 10{\cdot}62$), and the characteristics of the scattergram of attitude and comprehension were different from those of attitude and intelligence and knowledge. The t-tests showed the opposing group to have a significantly higher

comprehension score than the favouring or neutral groups without similar superiorities in intelligence and knowledge. The internal evidence, then, seems to suggest that there may be a genuine connexion between dislike and comprehension for this statement. If, however, the results for 71 N are considered in conjunction with those for the other twenty-three statements (i.e. as one significant correlation in twenty-four and a total of three significant *t*-test results in seventy-two comparisons for comprehension), both findings for 71 N can be regarded as chance results on statistical grounds. The fact that the other ambiguous statement, 67 S, did not give the same results as 71 N makes this a more likely explanation.

(iv) *Supplementary experiment*

Subjects in this experiment divided well into two attitude groups: right-wing and left-wing (i.e. mainly Conservative and Labour). These were treated as favouring or opposing groups according to whether the statement was of Conservative or Labour origin. There were no significant differences between the two groups on either the N.I.I.P. test or the political knowledge test.

Fisher's *t* was calculated for differences in comprehension on the Conservative statements 52 C, 36 F, 54 F, and on the Labour statements 63 C and 6 F. The differences were not significant. This confirms the results of the main experiment.

IX. Discussion

The experimental results do not support the view that comprehension of statements depends on the direction of the subjects' attitudes to the opinions expressed in them. This was the case for easy, average, and difficult sentences, despite the fact that the material was designed to, and did, arouse strong feelings on political matters. Many subjects expressed themselves most forcibly in the scripts. A qualitative analysis of the scripts corroborated the findings. Making sense of a sentence was seen to be like other attempts at solving problems. Subjects asked themselves questions and tried to answer these, or they tried out various meanings of each term separately. A number of subjects commented that they had used such a method. In other cases, it was clear from the scripts that this had been done. Correct interpretations were those by which the meanings given to each term fitted in with each other, and so made the sentence intelligible. Mistakes arose mainly when subjects got stuck with an interpretation of a term which—although possible—was not the correct one in the context. Instead of trying out another meaning, they retained their first interpretation, and either considered the statement to contain a contradiction, or they changed the remainder of the sentence to fit in with this sense of the word, or they ommited the term altogether while still giving some meaning to the sentence. The resultant misinterpretations could be classified as follows: (*a*) Substituting a familiar idea, (*b*) substituting a 'pet' idea, (*c*) believing an idea to be attacked which was, in fact, supported or not mentioned in the statement, (*d*) unwarranted elaborations, (*e*) omissions of essential parts of the statement. As in other experiments on attitudes and cognitive processes, the misinterpretations showed the subjects' bias. The occasion for the mistakes was, however, almost invariably a word or phrase which most subjects—whatever their political bias—had found difficult. Thus, the misinterpretations of a sentence by a Conservative, a Liberal, and a Socialist would be different in content, but the words which had been misconstrued or omitted were the same.

The subjects themselves also cited the construing of terms as their major difficulty.

They complained of (*a*) the vagueness of terms, (*b*) their level of generality, (*c*) the number of such vague or general terms in a sentence, (*d*) the number of subordinate clauses, (*e*) inability to give examples, and (*f*) illogicalities and contradictions in the statements.

The material suggests that mistakes mainly arose from lack of knowledge or skill in manipulating words. When an objective difficulty—too many vague terms or some other intellectual lacunae—prevented a correct interpretation, the subject fastened on whatever items were familiar and elaborated material which, to him, was associated with these. Thus, the *content* of mistakes frequently revealed emotionally associated material; the *number* and *kind of occasions* for mistakes—by which efficiency is measured—depended in the main on intellectual factors and seemed to be independent of direction of the subjects' political attitudes. In other words, a distinction must be made between the kind of occasions from which mistakes arise and the content of mistakes.

Caution should be used in generalizing these results, since the number of subjects was relatively small and they were drawn mainly from an undergraduate and graduate population. The findings suggest, however, that careful scrutiny is needed of statements which include social attitudes among the major determinants of efficiency of comprehension.

I am indebted to Prof. C. A. Mace for his help and guidance in supervising the investigation.

REFERENCES

BARTLETT, F. C. (1932). *Remembering.* London: Cambridge University Press.

BAUMGARTEN-TRAMER, F. (1941). Zur Psychologie des Verstehens und der Verständigung. *Gesundh. Wohlf.* no. 2, 61–4.

COOPER, E. & JAHODA, M. (1947). The evasion of propaganda. *J. Psychol.*, **23**, 15–25.

DAVIS, F. B. (1942). Fundamental factors of comprehension in reading. *Psychol. Bull.* **39**, 499–500.

DAVIS, F. B. (1944). Fundamental factors of comprehension in reading. *Psychometrika*, **9**, 195–97.

DOOB, L. W. (1947). The behavior of attitudes. *Psychol. Rev.* **54**, 135–56.

EDWARDS, A. L. (1941). Political frames of reference as a factor influencing recognition. *J. Abnorm. (Soc.) Psychol.* **36**, 34–50.

FISHER, R. A. (1950). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

FRIEDMANN, S. (1952). The influence of attitudes on the understanding of verbal material. Ph.D. Thesis, London 1952.

HIPPIUS, R. (1936). Gefühl und Denken, Ber. 15. *Kongr. Dtsch. Ges. Psychol.* Jena, 52–65.

HODGES, H. A. (1944). *Wilhelm Dilthey.* London: Routledge and Kegan Paul.

LEVINE, R. & MURPHY, G. (1943). The learning and forgetting of controversial material. *J. Abnorm. (Soc.) Psychol.* **38**, 507–17.

MARUM, O. (1936). Analyse des Verstehens und seiner Bedingungen. *Arch. ges. Psychol.* **95**, 99–115.

MASON, H. M. & GARRISON B. K. (1951). Intelligibility of spoken messages—liked and disliked. *J. Abnorm. Soc. Psychol.* **46**, 100–3.

RICHARDS, I. A. (1930). *Practical Criticism.* London.

RYLE, G. (1949). *The Concept of Mind.* London: Hutchinson.

SAKODA, J. M., COHEN, B. H. & BEALL, G. (1954). Tests of sigificance for a series of statistical tests. *Psychol. Bull.* **51**, 172–5.

SEELEMAN, V. (1940). The influence of attitudes upon remembering pictorial material. *Arch. Psychol.* no. 258.

SHERIF, M. & CANTRIL, H. (1945). The psychology of attitudes, I. *Psychol. Rev.* **52**, 295–319.

SHERIF, M. & CANTRIL, H. (1946). The psychology of attitudes, II. *Psychol. Rev.* **53**, 1–24.

VERNON, P. E. (1950). The intelligibility of broadcasts. Audience research department, British Broadcasting Corporation.

WILKINSON, B. A. (1951). A statistical consideration in psychological research. *Psychol. Bull.* **48**, 156–8.

(*Manuscript received* 29 *April* 1957)

TIME FOR READING AND MEMORY

By E. C. POULTON

Medical Research Council Applied Psychology Research Unit, Cambridge

This paper on silent reading is concerned with the rate of comprehension, selection while reading, and the rigidity of pacing while reading. Subjects studied 144 short statements under one of three degrees of pacing: (*a*) a time was set for studying the whole series of statements (whole condition), (*b*) each complete statement was presented in turn by itself for a set time (statement condition), and (*c*) part only of a single statement was presented at a time, and there was a period after each statement when no print was visible (segment condition). In different subconditions 1·0 min. was allowed for an average of every 293, 146, 73 or 37 words. In one variation of condition (*a*) the subject was instructed to pay attention only to the 18% of statements connected with geography (selection condition). In another variation, the time for study was fixed at 3·0 min., and the number of statements was reduced accordingly (shortened condition).

After reading, the subject had first to estimate from memory the percentage of statements which could be classified under certain specified topics. He was then given a test of memory for seventy-two of the statements. Half of them had to be completed (recall), and each of the remaining half had to be discriminated from a paired statement which resembled it in one of six standard ways (recognition). An identical test of memory for the remaining seventy-two statements was normally given unexpectedly 1 week later. A control group of subjects took these tests without having read the statements.

In the whole condition when 1·0 min. was allowed for every 293 or 146 words, most of the subjects reported that they had read practically all the statements once and only once. Reading at about 146 words per minute in this condition was found to give significantly better comprehension of the individual statements than reading at about 293 words per minute. In contrast, in this whole condition when 1·0 min. was allowed for every 73 or 37 words, the subjects reported that they had re-read some or all of the statements at least once. Between the rates of presentation of 37 and 293 words per minute, substance memory varied approximately as the logarithm of the time available per statement. After the 3·0 min. of study in the shortened condition, about the same number of statements could be remembered in substance or recognized whether 293 or only 73 words had had to be studied per minute. Selection tended to improve memory for the selected statements, and to reduce memory for the remainder. But when 293 words had to be studied per minute, only recognition soon after study benefited significantly from selection.

When 1·0 min. was available for every 146 words, the rigid pacing of the segment condition gave significantly better memory soon after study than the loose pacing of the whole condition. But when 1·0 min. was available for every 37 words, pacing had no effect upon memory. There were no significant differences between the degrees of pacing used in the statement and whole conditions.

I. Introduction

(i) *Speed of reading and comprehension.* The primary aim was to determine how comprehension is affected by altering the speed of reading to the greatest feasible extent. No conclusive experimental answer yet appears to have been given to this question. If the function relating speed of reading to comprehension corresponds to the typical learning curve relating time for study to memory (e.g. Woodworth, 1938, Fig. 6), speed of reading and comprehension would be inversely related. The only exception would occur at speeds of reading which gave the highest score on the test of comprehension which the particular reader was ever likely to obtain, when there would be no correlation. However, the reported correlations between speed of reading and comprehension have often been positive. And in describing the reading-training programme at Iowa State College, in which speed was used as a criterion of improvement but not comprehension, Lauer stated

(1936, p. 655) 'it is quite generally agreed that speed in reading is positively correlated with comprehension'. Some of this evidence will now be outlined, together with the reasons for disregarding it.

First, Tinker (1939) found that with all but the most difficult standardized reading tests, the faster undergraduate reader obtained a higher comprehension score than the slower reader. But this result appears to have been an artifact of the experimental method he used. For in these reading tests comprehension was assessed by multiple-choice questions answered by the reader as he went along. The reader, therefore, presumably adjusted his rate of reading to the difficulty of the questions. The net result was that mistakes were rarely made, except in the difficult passages where more than one interpretation was possible. Thus, with the easy passages the number of questions answered in a given time, which is a criterion of speed, determined the recorded score for comprehension. Speed of reading and comprehension were therefore found to be highly correlated. (Tinker actually took as his criterion of speed the time to complete the alternative version of each test (1940). But he stated that this criterion correlated almost perfectly with the number of questions answered in a given time in the original version.) A modification of this approach, which avoided the difficulty, was used by Stroud & Henderson (1943). They gave multiple-choice tests of recognition *after* the reading. They found a correlation of only about +0·1 between speed of reading and success in the tests. But even if the correlation had been higher, no valid deductions could have been made from this correlation to the relationship between speed and comprehension within a single individual. For the experimental design confounded the relationship between speed and comprehension with individual differences in previous knowledge (Poulton, 1957*a*), intelligence, and personality characteristics. The faster readers might have recognized more still if they had read rather more slowly.

A second line of evidence comes from the before and after scores of young adults trained to read faster. Generally a large and statistically significant improvement is found in the speed of reading, together with a smaller and often not significant improvement in comprehension as measured by multiple-choice recognition questions. But again no valid deductions which are relevant to our question can be made from data of this kind, for here the relationship between speed of reading and comprehension has been confounded with the effects of the training.

A third line of evidence is supplied by Perry & Whitlock (1948*b*, fig. on inside of back cover). They used, as an independent criterion of reading ability, the grades achieved in certain elementary history courses at Harvard University (Perry, 1955). They found that the same average grade was achieved by a moderately fast reader with moderate comprehension, a fast reader with poor comprehension, and a slow reader with good comprehension. Thus here the correlation was negative. Unfortunately, they did not ask the slow reader to read faster, nor the fast reader to read more slowly, and determine the changes in comprehension score. Thus these data also do not answer our question.

(ii) *Subsidiary aims.* In everyday life we cannot always ask ourselves: 'how long ought I to spend on this document?' We may have only a limited time available, and the question then becomes: 'how should I distribute my time?' Is it possible to learn as much by reading a part only of the document carefully, as by reading it all more hurriedly? A subsidiary aim was to answer this question, both when the part to be read was selected at random, and when the reader knew what topic to concentrate on.

A second subsidiary aim was to find out whether comprehension was affected by the degree of rigidity of pacing while reading. Excluding the tachistoscope, there are, in general, three degrees of pacing which have been used in training adults to read faster (Glock, 1949). The least rigid form is simply to hand the subject a passage, and to tell him to read it as quickly as he can. In the intermediate form of pacing, some kind of marker descends slowly down the page, and the subject has to keep ahead of it. In the most rigid form, only a part of a line of print can be seen at a time, as in the Harvard reading films (Perry & Whitlock, 1948*a*). Similar (though not identical) degrees of pacing were used in the present experiment.

(iii) *Reading material.* This consisted of a series of short disconnected statements, and was chosen for three reasons. First, comprehension of this type of material should be less adversely affected by increasing the speed of reading than comprehension of a developing argument. For comprehension of an argument is largely 'all or none'; insufficient time to understand any one of its stages means failure to understand the whole properly. Whereas with disconnected statements failure to understand one statement should not appreciably affect the chances of understanding the next statement. Thus, if there were any truth in the view that speed of reading is either positively correlated with comprehension, or at least unrelated to it, it would have a good chance of being confirmed with this type of material. Secondly, memory for each statement could be scored equally and objectively. The experimenter did not have to select the parts of a passage for which memory was to be tested. Thus there would be no question of a subject who selected what was worth remembering, much as the experimenter selected what was to be tested, doing better than a subject who made an unlucky selection. Thirdly, the material was the same as that used in an earlier experiment (Poulton, 1957*a*). Thus it would be possible to relate some of the results of the two experiments.

II. Methods

The same material was used as in a previous experiment (Poulton, 1957*a*). It is reproduced in full in the Appendix to Poulton (1957*b*). There were four lists, each of thirty-six statements. A rough classification of the statements according to topic is given in the first row of Table 6 below. The statements ranged in length from 3 to 9 words, with an average of 6·1 words. One of the longest was: 'The stern may be the rear of a steamer.' Each statement started with a title, e.g. 'The stern'; and contained two key words, e.g. 'rear' and 'steamer'. Some of the titles contained unfamiliar words, but all the key words were in everyday use. Half the statements in each list were at least approximately true, and the other half were to some degree false. The statements were typed one below the other on two paper tapes 6 in. wide. The typing was double spaced, twelve letter spaces to the inch. In all conditions except a shortened condition, one tape carried lists 1 and 2 without a break between them, and the other tape carried lists 3 and 4. In the shortened condition the number of statements on each tape was reduced (see below).

(i) *Reading.* The experiment was carried out in three parts: reading, immediate test of memory, and delayed test of memory. In the first part, each subject was tested individually. He studied the statements with the knowledge that he would be asked questions on them afterwards, although the nature of the questions was not revealed. There was a break of about 1·0 min. between the presentation of the two paper tapes. The eighteen subconditions of study are listed in Table 1. Each subcondition was performed by a separate group of eight subjects. The reading was always preceded by a suitable practice on eighteen statements, which allowed the subject to learn his particular procedure.

In a *statement* condition the statements were presented automatically in succession, through a horizontal slit in a screen. Each statement was shown once only for 10·0, 5·0, 2·5 or 1·25 sec., depending upon the subcondition. At the end of the permitted time, the next statement was moved smartly into position by a solenoid. The presentation rate of 1·25 sec. per statement required reading at about 293 words per minute.

In a *segment* condition a clear celluloid disk rotated behind the screen which contained the horizontal

Table 1. *Conditions of study*

Time for study		Time of exposure (sec.)					
Sec./statement	Words/min.	Segment	Statement	Whole	Selection	Shortened	Control
10·0	36·6 (37)	1·05	10·0	720	720	90	—
5·0	73·2 (73)	0·95	5·0	360	360	90	—
2·5	146·4 (146)	0·80	2·5	180	180	90	—
1·25	292·8 (293)	*	1·2	90	90	*	—
0·0	∞	—	—	—	—	—	0

Note. Each subcondition was performed by a separate group of eight subjects, except the control which was performed by forty-eight subjects.
* This subcondition was not used.

slit. The disk carried a black paper mask with a spiral window in it. Thus, the subject saw at a time only a segment of the typewritten line of a fixed size. This segment moved smoothly across the paper tape, so that new letters appeared on the right as letters disappeared on the left. As soon as the window moved off the tape at the right-hand edge, and before it appeared again on the left, the next statement was placed in position. The apparatus will be described in greater detail in a subsequent paper. In the present experiment when 10·0 sec. was allowed per statement, the size of the visible segment was nine letter spaces. A new set of letter spaces appeared every 1·05 sec. Because the average length of a statement was only 2·5 in., while the width of the paper tape was 6·0 in., no letters were visible on average for 5·6 sec. out of the 10 sec. allowed per statement. During the time that the subject was actually reading, he had therefore to read at an average speed of 84 words per minute. At 5·0 sec. per statement, nineteen letter spaces could be seen at a time, a new set appeared every 0·95 sec., and no letters were visible on average for exactly half the time. The average speed of actual reading was therefore 146 words per minute. At 2·5 sec. per statement, thirty-seven letter spaces could be seen at a time, a new set appeared every 0·80 sec., and no letters were visible on average for 1·05 sec. out of the 2·5 sec. available, so that the average speed of actual reading was 252 words per minute. (No presentation rate of 1·25 sec. per statement was used, because in a pilot experiment subjects reported that they could not see each statement for long enough to read it, even with a moving window the full width of the paper tape.)

In a *whole* condition the subject was handed a complete paper tape at a time, and told to study it for the total time permitted for the seventy-two statements. The experimenter called out when one-, two-, and three-quarters of the time had passed, and the ends of the quarters were marked on the tape. The subject was told that at a quarter time he must leave the first quarter of the statements, if he had not done so already; at half time he must leave the second quarter, and so on. If he was ahead of schedule, he was told to employ whatever time was left over in additional study. When time was up, he had to report how many times he had read each statement. The same four rates of presentation were used as in the statement condition.

In a *selection* condition the subject had to study only the geographical statements. These were defined as statements 'concerned with other countries or parts of this country'. The procedure was otherwise the same as in the whole condition.

In a *shortened* condition the subject was allowed only 1·5 min. for study in each half of the period of reading. Special shortened paper tapes had therefore to be prepared. No presentation rate of 1·25 sec. per statement was used, since this subcondition would have been identical with the corresponding whole subcondition. When 2·5 sec. were available per statement, four subjects were given lists 1 and 2, respectively, for the two halves of the period of reading. The other four subjects were given lists 3 and 4. The order of studying the lists was always counterbalanced for different subjects in all subconditions. When 5·0 sec. were available per statement, two subjects were given only the first halves of lists 1 and 2, respectively, for the two halves of the period of reading, and two subjects were given the second halves of lists 1 and 2. Lists 3 and 4 were divided in similar fashion between the other four subjects. When 10 sec. were available per statement, half the statements in each half list were not used at all. The design was otherwise similar to that used at 5 sec. per statement. The blocks of statements selected and rejected were equated for difficulty as judged by the results of the other experimental conditions.

(ii) *Tests of memory.* The second part of the experiment followed directly after the first. The subject was taken to another room, where one or more other subjects might already be working. First, he had to fill in his personal details on a form. Secondly, he had to estimate the percentage of the statements which fell

into the following five categories: '(*a*) The navy and life at sea. (*b*) History, both old and contemporary. (*c*) Geography (places and people in other countries, and in other parts of this country). (*d*) Science, including physics, engineering, biology and medicine. (*e*) Other subjects.' To help him, he was given two examples of each type, taken from the practice statements (see Appendix to Poulton, 1957*b*). The order of the first four topics on the form was different for different groups of subjects.

Then came a test of recall, identical with that given in the previous experiment (Poulton, 1957*a*). The titles of one list of statements were presented on an otherwise blank sheet, and the subject had to complete as many of the statements as he could. In the shortened condition, the list was selected from the pair of lists of which the subject had read at least some of the statements, and he was asked to guess the answers to any statements which he had not seen previously.

Finally came a test of recognition, which was again identical with that of the previous experiment. The statements of a second list were presented in full. (In the shortened condition this list was the second of the pair of lists of which at least some of the statements had been read.) Each statement had a similar statement typed either above or below it, from which it had to be distinguished. The alternative version had the same title as the original. Half the alternative versions had one different key word, and the other half had both key words different. Each of these subgroups was divided into three, according to meaning. The meaning of the alternative was either more or less the same as the original, in some sense the opposite, or as far as possible unrelated (see Appendix to Poulton, 1957*b*). If the subject was not sure which of a pair to select, he was made to guess. Unlimited time was available for all these tests of memory.

A delayed test of memory was given 1 week later, except in the shortened condition. It came as a surprise to all subjects. Six subjects were tested at a time. The test employed the other two of the four lists of statements read. Recall of the individual statements of one list was tested first, followed by recognition of the statements of the other list. The two tests were similar to the corresponding tests of the previous week.

(iii) *Balanced design.* The order of reading the two paper tapes, the times allowed for reading, the order on the form of the four specific topics in the first test of memory, and whether the lists of statements had to be recalled or recognized, were all balanced in a multifactorial design for each condition. Whether memory for a list was tested immediately or after a week was also balanced in the multifactorial design, except in the shortened condition which had no delayed test of memory.

(iv) *Control group.* A group of forty-eight subjects had to take the immediate tests of memory without having read the statements. The procedure was similar to that used for the control group in the previous experiment (Poulton, 1957*a*) but, in addition, the subjects first made estimates of the percentage of statements which could be classified under each of the five headings. The order on the form of the four specific topics was again balanced for different subgroups. On each list of statements twelve subjects then took the test of 'recall'; another twelve subjects took the test of 'recognition', after a test of 'recall' on one of the other lists.

(v) *Subjects.* The 192 subjects were all naval ratings drawn from the same population as the subjects of the previous experiment (Poulton, 1957*a*). Their ages ranged from 17 to 36, with a median of 21. Their scores on intelligence test AH 4 (Heim, 1955) ranged from 31 to 121, with a median of 70. Most had left school at about 15, and had been apprentices or had taken miscellaneous jobs before joining the navy. The experimental subgroups were selected to be of approximately similar composition.

(vi) *Scoring.* The same methods of scoring were used as in the previous experiment (Poulton, 1957*a*). Correct literal recall of a statement required memory for the two key words, but not the exact verbal structure of the original. Correct substance recall required memory for the meaning of at least one of the two key words, expressed in a form which did not alter too grossly the general sense of the other key word.

In order to assess the effect upon memory for geographical and non-geographical statements of selective study of the geographical statements (Table 5), sixteen of the 144 statements were classified as unambiguously geographical (see Appendix to Poulton, 1957*b*). Each of these statements contained the name of a place or country. Another ninety-nine statements were classified as definitely non-geographical. These two forms of statement were divided more or less equally between the four lists. The remaining twenty-nine statements were not considered in making this comparison, as they could be taken to be borderline cases.

In order to assess the effect upon the memory scores of the possible interaction of time for study with previous knowledge, three sets of six statements each were selected from each list (see Appendix to Poulton, 1957*b*). The criteria for selection were respectively: (*a*) the six statements in each list which had the highest indices of general knowledge in the previous experiment (Poulton, 1957*a*); (*b*) the six statements with an unfamiliar word in the title which received the greatest number of 'no idea' ratings in the

previous experiment; and (*c*) the six statements in each list without an unfamiliar word in the title which had the lowest indices of general knowledge.

(vii) *Calculations.* As in the previous paper (Poulton, 1957*a*), a recognition score was computed by doubling the number of errors in the test of recognition, and subtracting this product from the number of pairs of statements presented. Wherever possible the significance of the difference of each pair of means was assessed separately (*a*) against the variability between subjects, and (*b*) against the variability between statements. The level of the less significant of the two tests was taken. This test was generally (*a*). Two-tailed tests were used except where stated. As compared with more conventional methods, this method tends to underestimate significance. For the data in Tables 3 and 6, only tests of type (*a*) could be used.

A table was sometimes found to contain few significant differences. In this case the total number of comparisons made was taken into account in assessing the importance of the results, after the method described by Sakoda, Cohen & Beall (1954). If this overall assessment was not significant, the individual differences which were found to be significant have not always been mentioned in the results, unless they were to have been expected on other grounds.

The significance of the difference in position between the two top curves in Fig. 1 was determined by comparing each of the three corresponding points which did not coincide, using *t*-tests. The three *P*'s were then combined by Fisher's χ^2-method (Fisher, 1948, pp. 99–101). The lowest point was common to both curves, and so could not be used.

The standard deviations in Table 3 were calculated from the two verbal reports of each subject, one for each half of the period of reading. Thus, they cannot be used to estimate the significance of the differences between conditions, since they include variability within subjects as well as the variability between subjects.

The scores in the lower half of Table 4 (and of the lower broken curve in Fig. 1) were computed from the scores in the upper half of the table (or of the upper broken curve), by allowing the same proportion of correct guesses on unseen statements as was achieved by the forty-eight control subjects. The data from the shortened condition could not be used as they stood, because the group of subjects who studied at 146 words per minute (2·5 sec. per statement) did not do any guessing; they had simply to recall the statements of one of the two lists which they had read, and to recognize the statements of the other list (see above). The guessing of the two groups of subjects who studied at 73 and 37 words per minute, respectively, was very similar to the guessing of the control group. The significances of the differences in the lower half of Table 4 were calculated from the actual scores in the shortened condition.

The significance of the influence of the order of the topics on the form upon the proportion of statements classified under each topic (Result (vii)) was estimated by analysis of variance.

III. RESULTS

(i) *Time for study and memory.* The effect of the time allowed for study of the 144 statements upon memory is shown in Table 2 for the three degrees of pacing. The scores for the guessing of the control subjects are given in the first column. In the last column are the scores from a condition of the previous experiment (Poulton, 1957*a*), in which one statement was heard every 10 sec. By all the criteria of memory used, the amount remembered increased significantly as the time allowed per statement was increased. For all degrees of pacing combined, memory after a week was still better than guesswork (no study), except for recall after 293 words had been presented per min. Substance memory was actually a little below the level of chance in this latter case, because the subject would not always guess when he had forgotten a statement.

Fig. 1 shows graphically the relationship between substance memory and the number of words to be studied per minute, in both the whole and the shortened conditions. If we consider only memory soon after study for the statements actually studied, in both conditions (between 293 and 37 words to be studied per minute) substance memory varied inversely approximately as the logarithm of the number of words to be studied per minute. However, the complete theoretical curves are presumably S-shaped, and this tendency is apparent in Fig. 1. (The function for the shortened condition lies above the

Table 2. *Effect upon memory for* 144 *statements of time for study and degree of pacing*

Criterion of memory	No study (control) (%)	Words to be studied/min. 293 Statement (%)	293 Whole (%)	146 Segment (%)	146 Statement (%)	146 Whole (%)
Memory soon after study						
Two key words reproduced	0·9*	10	13	27†	14	17‡
Meaning more or less reproduced	8·7*	23	22	48†	37‡	32‡
Recognition score	−1*	20	22	66	67‡	44‡¶
Memory after 1 week						
Two key words reproduced	0·9*	1·4	1·0	3·1	3·1	3·5‡
Meaning more or less reproduced	8·7*	7·3	7·3	15	10	13‡
Recognition score	−1*	2	22††	38	29‡	28

Criterion of memory	No study (control) (%)	Words to be studied/min. 73 Segment (%)	73 Statement (%)	73 Whole (%)	37 Segment (%)	37 Statement (%)	37 Whole (%)	37 Listening (%)
Memory soon after study								
Two key words reproduced	0·9*	35	32§	26§	32	32	28	45‖
Meaning more or less reproduced	8·7*	59	52§	50§	55	60	56	71‖
Recognition score	−1*	74	74	77§	80**	83	69	92‖
Memory after 1 week								
Two key words reproduced	0·9*	5·2	8·0	7·3§	3·8	5·2	6·3	10
Meaning more or less reproduced	8·7*	18	18§	19§	18	18	16	25
Recognition score	−1*	29	54§‡‡	43§	42	59	57	64

Note. The segment condition was the most rigidly paced, and the whole condition was the least rigidly paced. A negative recognition score is below the level of chance.

* For all degrees of pacing combined, no study different from all rates of study at the 0·05 level of confidence with a one-tailed test or better, except 293 words per minute on the two measures of recall after 1 week.
† Segment different from both statement and whole at the 0·05 level or better.
‡ 293–146, $P = 0{\cdot}05$ with a one-tailed test or better.
§ 146–73, $P = 0{\cdot}05$ with a one-tailed test or better.
‖ Listening—statement, $P = 0{\cdot}05$ or better.
¶ Whole different from both segment and statement at the 0·01 level.
** 146–37, $P = 0{\cdot}05$.
†† Statement—whole, $P = 0{\cdot}05$.
‡‡ Segment—statement, $P = 0{\cdot}05$.

function for the whole condition, but the difference was not significant. The difference was presumably due to the fewer statements which had to be memorized in the shortened condition at the slower rates of study. Thus when only 37 words had to be studied per min., there were only eighteen statements to be memorized in the shortened condition, as compared with the full 144 statements in the whole condition.)

(ii) *Reported speed of reading.* The mean numbers of times the lists were reported to have been gone over in the loosely paced conditions are shown in Table 3. The standard deviations of the means for the two halves of each period of reading are also given (see Calculations). Except when 293 words had to be studied per minute, the lists were gone over least frequently in the whole condition, and most frequently in the selection condition.

In the whole condition the average reported speeds of reading were thus 299, 170, 126 and 133 words per minute respectively, for the four subconditions. When 293 words had to be studied per minute in the whole condition, no subject had time over at the end in which to re-read more than a very few of the statements, and two of the eight subjects

failed to read quite all of them. (The memory scores of these two subjects were about the same as the scores of the remainder of the subjects.) Thus in this subcondition the time allowed for study in words per minute corresponded fairly closely to the reported average speed of reading of each subject.

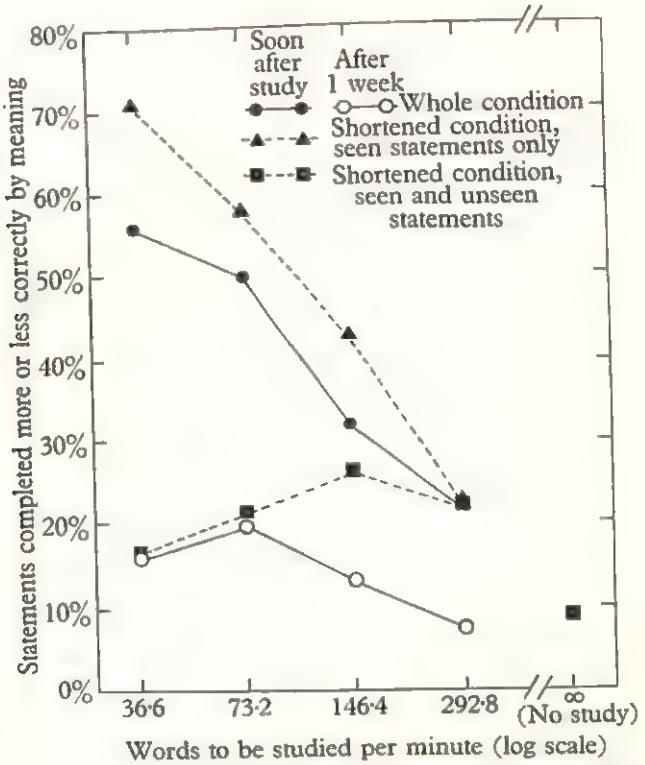


Fig. 1. The relationship between words to be studied per minute and substance memory in the whole and shortened conditions. Each point on a curve is based upon data from a separate group of eight subjects. The points on the lower broken curve were computed from the corresponding points on the upper broken curve, by allowing 8·7 % correct responses for unstudied statements. This is the value for no study which was given by the control group of forty-eight subjects.

Table 3. *Number of times lists were gone over*

	Words to be studied/min.							
	293		146		73		37	
Condition	Mean	S.D.	Mean*	S.D.	Mean†	S.D.	Mean	S.D.
Whole	1·02	0·048	1·16	0·23	1·72	0·56	3·63	1·01
Shortened	—	—	1·44	0·39	2·41	0·71	4·16	1·38
Selection	0·97	0·042	1·70	0·64	2·73	0·43	4·20	1·74

* Whole different from selection at the 0·05 level of confidence.
† Whole different from shortened at the 0·05 level, and from selection at the 0·001 level.

In the whole condition when 146 words had to be studied per minute, all the subjects finished reading the statements, and six of the subjects had time to re-read only a very few of them. The reported average speed of reading of these six subjects was about 153 words per minute. The two remaining subjects re-read just under and just over half the statements, respectively, but their memory scores were about average. Thus, the memory

scores for this subcondition can be taken as reasonably representative of subjects actually reading at about the required rate of study (146 words per minute).

In contrast, in the whole condition when seventy-three words or less had to be studied per minute, the time allowed for study no longer corresponded to the average reported speed of reading. When these long times were available, the rigid pacing of the segment condition was the only method of ensuring that each statement was read once and only once.

(iii) *Reported speed of reading and memory.* We have just seen that, in the whole condition when 293 or 146 words had to be studied per minute, either the reported speed of reading in words per minute corresponded fairly closely to the time allowed for study, or at least the memory scores were reasonably representative of subjects reading at about the set speed. Table 2 shows that, in this condition, memory was significantly better when 146 words had to be studied per minute, than when 293 words had to be studied per minute, by every criterion except delayed recognition. Thus, we may conclude that the amount remembered increased significantly when the average reported speed of reading decreased from about 293 to about 146 words per minute.

The effect of reducing the average speed of reading below about 146 words per minute could only be determined in the segment condition. The rigid pacing in this condition ensured that the subject read each statement once only at a controlled speed. The range of controlled reading speeds ran from an average of 252 words per minute at the presentation rate of 146 words per minute, to 84 words per minute at the presentation rate of 37 words per minute (see Method). Table 2 shows that, in this segment condition, recognition soon after study was significantly better after reading at the slow speed than after reading at the fast speed. The other criteria of memory showed similar trends, but they were not significant statistically.

(iv) *Degree of pacing and memory.* Table 2 also shows the effect upon memory of the degree of pacing imposed during the reading. Soon after 146 words had been presented per minute, the rigid pacing of the segment condition gave significantly better memory scores than the loose pacing of the whole condition. Yet the distributions of scores on the intelligence test AH4 (Heim, 1955) were almost identical for the two groups of subjects. The segment condition also gave better recall than the whole condition soon after 73 words had been presented per minute, but the differences were not significant.

There were no obvious differences in memory between the three degrees of pacing soon after only 37 words had been presented per minute. The only significant difference soon after study between the statement and whole conditions was in recognition after 146 words had been presented per minute. This difference could have been due to chance, since in twelve comparisons one significance at the 0·01 level has a probability of occurrence greater than 0·1 (Sakoda *et al.* 1954, fig. 2). Similarly, the two significant differences in memory after a week between the three degrees of pacing could have been due to chance. For in twenty-one comparisons, two significances at the 0·05 level have a probability of occurrence of about 0·25 (Sakoda *et al.* 1954, fig. 1).

(v) *Distribution of study and memory.* Table 4 shows the effect upon memory soon after study in the shortened condition of the amount to be studied in the two periods of 1·5 min. each. The upper half of the table gives the scores for the statements actually studied. As in Table 2, memory for the statements studied increased significantly as the time allowed per statement was increased. The lower half of Table 4 gives computed scores for all 144

statements, whether studied or not (see Calculations). Both for substance memory and for recognition, the scores for the whole 144 statements were about the same whether an attempt had been made to study all 144 statements (293 words to be studied per minute), or whether only one-quarter of the statements had been studied in the time (73 words to be studied per minute). There was a suggestion that it may actually have been better to have concentrated on only half the statements (146 words to be studied per minute), than to have attempted to study them all. However, when recall of the actual words was the criterion, memory soon after study was rather better when all the statements had been looked at.

Table 4. *Effect of amount to be studied in 3 min. upon memory soon afterwards*

		Words to be studied/min.			
Criterion of memory	No study (control) (%)	293 (%)	146 (%)	73 (%)	37 (%)
Statements actually studied					
Two key words reproduced	—	13*	18	30	39
Meaning more or less reproduced	—	22	43†	58	71
Recognition score	—	22	60‡	85	84
Computed scores for all statements (seen and unseen)					
Two key words reproduced	0·9	13	10	8·1	5·6
Meaning more or less reproduced	8·7	22	26	21	17§
Recognition score	−1	22	30	21	10§

Note. A negative recognition score is below the level of chance.
* 293 different from 73 at the 0·05 level of confidence.
† 146 different from 293 at the 0·01 level and from 37 at the 0·05 level.
‡ 146 different from both 293 and 73 at the 0·01 level.
§ 146 different from 37 at the 0·02 level or better.

(vi) *Selective study and memory.* Table 5 compares the selection condition with the whole condition on memory soon after study. Soon after 293 words had had to be studied per minute, it was only recognition of the statements to be selected which gave a significant difference in favour of selection. Recall of the statements to be selected was no better in the selection condition than in the whole condition. And on the statements not to be selected, selection resulted in significantly worse recall, and somewhat worse recognition, than simple study. In contrast, soon after only 37 words had had to be studied per minute, selection produced considerably better recall of the statements to be selected than full study, and not much worse memory of the statements not to be selected.

After a week, there were only two differences in memory for the statements to be selected which were significant at the 0·05 level. Both differences favoured selection, but neither occurred after 293 words had to be studied per minute. They could have been due to chance, since for twelve comparisons the combined $P = 0{\cdot}12$ (Sakoda *et al.* 1954, fig. 1). There were also two differences in memory after a week for the statements not to be selected which were significant at the 0·02 level. Both differences favoured the whole condition.

(vii) *Classification from memory soon after study.* The mean estimates of the percentage of statements classified under the different headings are shown in Table 6. The complete set of differences between the various times allowed for study, and between the three degrees of pacing, could have been due to chance (Sakoda *et al.* 1954). The data for the four different times allowed for study, and for the statement and whole conditions, have

Table 5. *Effect of selective study upon memory soon afterwards*

Criterion of memory	No study (control) (%)	Words to be studied/min. 293 Selection (%)	293 Whole (%)	146 Selection (%)	146 Whole (%)	73 Selection (%)	73 Whole (%)	37 Selection (%)	37 Whole (%)
Statements to be selected									
Two key words reproduced	2·6	13	13	28	19	56	25*	75	25†
Meaning more or less reproduced	10	19	31	50	28*	63	44	81	50‡
Recognition score	−11	75	19‡	62	62	100	87	94	75
Statements not to be selected									
Two key words reproduced	0·6	1·5	14†	4·5	19†	21	26	18	29
Meaning more or less reproduced	7·2	5·1	21†	16	34†	35	51	41	56
Recognition score	−2	12	25	24	41	41	74*	66	70

Note. A negative recognition score is below the level of chance.
* Selection—whole $P=0{\cdot}05$.
† Selection—whole $P=0{\cdot}01$ or better.
‡ Selection—whole $P=0{\cdot}02$.

therefore been pooled. The data for the segment and shortened conditions have not been included, because they were not performed at the fastest rate of presentation (see Table 1). The approximate actual percentages in the top row of the table were calculated from the experimenter's own classification (see Appendix to Poulton, 1957*b*). The bottom row gives the theoretical estimates if all five categories had been equally probable. For each heading in Table 6, the value for the combined statement and whole condition lies on the same side of twenty as the actual percentage. The actual percentages for navy and other topics are the two furthest from twenty, and in these two cases the differences between the combined statement and whole conditions and the control condition (no study) are highly significant. Selection gave similar estimates to the combined statement and whole conditions.

Table 6. *Classification of statement soon after study*

Condition or basis of assessment	Mean percentage judged to be concerned with: Navy*	History	Geography	Science†	Other topics*
Approximate actual percentage	7	17	18	26	32
Combined statement and whole conditions	11·6	18·7	19·8	22·8	27·1
Selection condition	10·7	20·1	22·2	19·8	27·2
Control condition (no study)‡	18·3	18·2	21·6	25·8	16·1
Percentage if all categories were equally probable	20	20	20	20	20

* Control different from both combined statement and whole, and from selection, at the 0·001 level of confidence.
† Control different from selection at the 0·02 level.
‡ Science different from all others at the 0·05 level or better; geography different from all others at the 0·05 level with one-tailed test or better.

Two differences which are not shown in Table 6 are in keeping with the results in Table 2. First, for the combined statement and whole conditions, the mean for navy was 13·0 after 293 or 146 words had to be studied per minute, as compared with a mean of 10·2 after 76 or 37 words had to be studied per minute ($P=0{\cdot}05$). And secondly, after 146 words had to be studied per minute, the mean for navy was 13·4 in the statement condition, as compared with 7·9 in the segment condition ($P=0{\cdot}05$ with a one-tailed test). The navy was the only specified topic upon which significant differences were to have been

expected, for Table 6 shows that only with this topic was there a large discrepancy between the actual proportion and the guessed estimates of the control subjects.

The control subjects had to make their estimates without any knowledge of the nature of the statements. Table 6 shows that they tended to be most generous in their estimates of Science and Geography, which were the two categories with the longest descriptions (see Method). Both the control and the experimental subjects tended to give extra-generous treatment to the first two topics listed on the form, whichever they were, as compared with the other two specific topics listed ($P=0{\cdot}05$ or better).

(viii) *Other points.* We also investigated the effect of previous knowledge upon the relationship of the time allowed for study to memory for the individual statements. The data from the statement and whole conditions were combined, since there were no significant differences between these two conditions which could not be attributed to chance (Result (iv)). The difference between the rates of presentation of 293 and 37 words per minute was found to be significant ($P=0{\cdot}05$ a one-tailed test or better) by all the criteria of memory soon after study, and also by the criterion of delayed recognition, for three degrees of previous knowledge. These three degrees were represented by: (*a*) statements about whose truth the subject was likely to have been certain and correct (see Scoring); (*b*) statements which contained an unfamiliar word in the title; and (*c*) other statements about whose correctness he was likely to have had no idea. The statements in category (*a*) tended to be remembered best by all the criteria of memory, and to show the greatest rate of improvement in memory soon after study as a longer time was allowed for study. The statements in category (*b*) tended to be remembered least well, and to show the smallest rate of improvement in memory soon after study as the time for study was increased.

For the statement and whole conditions combined, the difference between the rates of presentation of 293 and 37 words per minute was also significant for both the true and the false statements by all but one of the criteria of memory used. The only exception was delayed memory of the actual words used in the false statements, and in this case the difference between 293 and 73 words per minute was significant. Soon after study there was no difference between memory for the true and the false statements. But after a week the meanings of the true statements were recalled rather better than the meanings of the false statements ($P=0{\cdot}02$), and the true statements also obtained higher scores on recognition ($P=0{\cdot}05$). These two results in favour of the true statements were probably due to successful guessing, for identical but more marked effects were shown by the control subjects who had not read the statements ($P=0{\cdot}002$ or better). In addition, in sixty-four out of the sixty-five instances in which the meaning of the completed statement of a control subject was in some sense the opposite to the meaning of the original statement, the original meaning was false ($P=0{\cdot}001$). Thus when the control subject guessed, he tended to make up or select a true statement. The experimental subject probably did the same.

Table 2 shows that there were significant differences in memory soon after study between the statement condition and the listening of the previous experiment (Poulton, 1957*a*). Yet in both cases the average rate of presentation was 37 words per minute (one statement every 10 sec.). These differences were associated with a number of major changes in procedure. The data from the listening experiment were only included for the sake of completeness.

We calculated the Tau correlations by subjects between the scores on intelligence test AH4 (Heim, 1955) and the various criteria of memory. Soon after study the correlations tended to be small and positive or negative (in one case significantly negative) after 293 or 146 words had been presented per minute, but about 0·4 after 73 or 37 words had been presented per minute. For delayed memory the correlations tended to be small, and all but one were positive.

IV. Discussion

(i) *Faster reading gave poorer comprehension.* The data in Table 2 (and Fig. 1) for substance memory in the whole condition when 293 and 146 words had to be studied per minute, show the relationship between the approximate speed of reading and comprehension. For (*a*) substance memory soon after reading is a criterion of comprehension; (*b*) the subjective reports of re-reading in this condition showed that at these two faster speeds most of the subjects had time to read practically all the statements once and only once (Result (ii)); and (*c*) the memory scores of the subjects with whom this relationship held least closely, were about the same as the scores of the remaining subjects (Result (ii)). We may therefore conclude from Table 2 that reading at about 293 words per minute gave significantly poorer comprehension than reading at about 146 words per minute.

With longer times for study in the whole condition, Table 3 shows that the subject simply re-read some or all of the statements. The average speed of reading was not reduced below about 130 words per minute in this condition (Result (ii)). The same was probably true in the statement condition, although subjective reports are not available. Thus we cannot use the data from these conditions to determine the effect upon comprehension of reducing the speed of reading below about 146 words per minute. However, the rigid pacing of the segment condition forced the subject to read each statement once only. When 37 words were presented per minute in this condition, the subject had to read at a speed of only about 84 words per minute (see Method (i)). Table 2 shows that recognition soon after study in this subcondition was significantly better than when 146 words were presented per minute, which meant actual reading at about 252 words per minute. Similar trends were shown by the two measures of recall soon after study, but they were not significant statistically.

However, the test of recognition was not as adequate a criterion of comprehension as was recall of the meaning of the statements. As pointed out in the previous paper (see Poulton, 1957*a*, Method (vii)), none of the six forms of the alternative versions of the statements presented the subject with a pure test of recognition of meaning, unconfounded by changes in the words used to carry the meaning. For in order to change the meanings of the alternative versions, some change in wording was necessary. The only way to have got round this difficulty would have been to have changed the wording of both the 'original' and the alternative versions. This would have forced the subject to select the statements with the meanings which were closer to the originals, without being able to recognize specific words from the original statements.

The relationship between speed of reading and comprehension, and the other relationships which we shall soon discuss, were all determined for a particular range of levels of reading skill, which was probably not a very high one. Thus, an extensive course of training before the experiment would presumably have improved the scores of most of the subjects. However, it seems unlikely that training would have altered the *direction* of these relationships. (This could be checked on suitably-trained subjects.)

(ii) *When time was short, careful study was not a disadvantage.* The lower broken curve in Fig. 1 shows that, in the shortened condition when only 3 min. were available for study, the total amount of substance memory soon afterwards was about the same whether all 144 statements had to be studied (293 words to be studied per minute), or whether only thirty-six statements had had to be studied (73 words to be studied per minute). The lower half of Table 4 shows that the same relationship held for recognition. It was only the total amount of memory of the specific words presented which increased consistently as more statements had to be studied in the time.

These results suggest that there is a limit to the amount of material which can be understood in a given time. If we proceed faster than this, we may be able to recall slightly more words; but we shall not be able to recall more of the meaning, nor to recognize more on a test of recognition similar to the one used here. To use the terminology of information theory (Quastler, 1955), the rate of coding the information in reading appears to have an upper limit, which is reached before the upper limit on the rate at which information can be stored in an uncoded form.

The results in the lower half of Table 4 were obtained by simply allocating different blocks of statements to different subjects, and then computing their scores for all the statements, whether they had studied them or not (see Calculations). If certain statements were more important than others, and the subjects knew how to select these statements, careful study of these statements should prove a positive advantage. This seems pretty obvious, and the top half of Table 5 confirms this impression. However, when 293 words had to be studied per minute, it was only recognition soon after study which benefited significantly from selection. Recall soon after study and delayed memory only benefited significantly when fewer words had to be studied per minute. This result is in line with the findings of the previous paper (see Poulton, 1957*a*, Discussion (vi)): recognition tends to become a more sensitive measure as conditions become more difficult.

If this result can be generalized to everyday reading, it suggests that reading selectively at high speed will only be an advantage over reading unselectively, if soon afterwards the reader returns to the information which he has selected (as, for example, in taking notes after reading). In the selection condition when 146 words or less had to be studied per minute, most subjects reported that they did in fact re-read at least once some or all of the selected statements (Table 3). But trainees on courses designed to improve speed and comprehension in reading are not necessarily warned to go over the important points a second time. On these courses a test of recognition normally follows immediately after the reading of a prose passage, and thus takes the place of reference back to the text (Perry & Whitlock, 1948*a*, *b*). If the present view is correct, the trainees ought to be instructed to substitute in their normal reading some kind of recapitulation for the test of recognition.

(This discussion would not be complete without pointing out the advantage to the reader who has not much time of clear short summaries and meaningful subtitles, for example, Experiment on —, *not* Experiment I. For this reduces the amount of essential reading and selection. A number of general points of this kind for the authors of papers have been made by Howie (1951)).

(iii) *Rigid pacing sometimes helped.* Table 2 shows that when 146 words had to be studied per minute, memory soon after study was significantly better with the rigid pacing of the segment condition, than with the loose pacing of the whole condition. This effect is un-

likely to have been due to a chance difference between the groups of subjects allocated to the two conditions, for two reasons. First, the distributions of scores on intelligence test AH4 (Heim, 1955) were almost identical for the two groups (Result (iv)). Secondly, recall showed a similar trend soon after 73 words had been presented per minute, although this trend was smaller and not significant (Table 2).

Two other explanations of this effect can also be ruled out: (*a*) that the automatic reading pacer used in the segment condition simply added novelty to the experiment, and thus motivated the subject more highly; and (*b*) that the whole condition confronted the subject with the additional task of ensuring that he reached the end of the list of statements on time. For first, when 146 words had to be studied per minute, the segment condition proved significantly better than the statement condition. This was the case both for recall of the individual statements soon after study (Table 2), and for the estimation of the proportion of statements 'connected with the navy and life at sea' (Result (vii)). Yet (*a*) both these two conditions involved using the automatic reading pacer; and (*b*) neither of them involved the additional task. Secondly, soon after study there were no significant differences which could not be attributed to chance between the whole and statement conditions (Result (iv)). Yet (*a*) the statement condition employed the reading pacer, while the whole condition did not; and (*b*) the whole condition involved the additional task, while the statement condition did not.

The difference between the segment and whole conditions when 146 words had to be studied per minute thus suggests that rigid pacing itself helped. In the segment condition at this rate of presentation, the subject had to read at about 252 words per minute for about 1·5 sec., and then had a break of about 1·0 sec. before the next statement was presented (see Method (i)). During this time he could memorize the statement which he had just read. This procedure appears to have been more effective than reading more or less continuously at about 146 words per minute, which was the procedure apparently used by most of the subjects in the whole condition (Result (ii)), and probably also in the statement condition.

Table 2 also shows that when only 37 words had to be studied per minute, the nature of the pacing had no effect upon memory. Yet in the segment condition each statement had to be read once only at an average speed of 84 words per minute (see Method (i)), and the reading was followed by an average interval of 5·6 sec. during which no print could be seen. Whereas in the whole condition each statement was read on average about three and a half times (Table 3), and the same was probably true in the statement condition. Clearly at a rate of presentation as slow as 37 words per minute, memory must have depended principally upon some factor related more closely to the overall time available than to the method of study.

This work was carried out in the Applied Psychology Research Unit under the direction of Dr N. H. Mackworth. I am also indebted to Sir Frederic Bartlett, F.R.S., and to Miss K. D. Napier for their interest. Mr M. Stone gave advice on statistical problems and Dr J. F. Mackworth on presentation. The Royal Navy supplied the subjects, and also the assistance of a sickberth attendant when I had to be in two places at once. Financial support from the Medical Research Council is also gratefully acknowledged.

REFERENCES

FISHER, R. A. (1948). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.
GLOCK, M. D. (1949). The effect upon eye-movements and reading rate at the college level of three methods of training. *J. Educ. Psychol.* **40**, 93–106.
HEIM, A. W. (1955). *Manual for Group Test of General Intelligence AH* 4. London: National Foundation for Educational Research.
HOWIE, J. W. (1951). On writing to be read. *Lancet*, **261**, 320–1.
LAUER, A. R. (1936). An experimental study of the improvement in reading by college students. *J. Educ. Psychol.* **27**, 655–62.
PERRY, W. C., JR. (1955). Personal communication.
PERRY, W. C., JR. & WHITLOCK, C. P. (1948*a*). *Harvard University Reading Films, Series Two.* Cambridge, Mass: Harvard University Press.
PERRY, W. C., JR. & WHITLOCK, C. P. (1948*b*). *Selections for Improving Speed of Comprehension.* Cambridge, Mass.: Harvard University Press.
POULTON, E. C. (1957*a*). Previous knowledge and memory. *Brit. J. Psychol.* **48**, 259–70.
POULTON, E. C. (1957*b*). Time for reading and memory. Cambridge: *Appl. Psychol. Unit Rep.* no. 302. (Available from the author.)
QUASTLER, H. (Ed.) (1955). *Information Theory in Psychology.* Glencoe, Ill.: Free Press.
SAKODA, J. M., COHEN, B. H. & BEALL, G. (1954). Tests of significance for a series of statistical tests. *Psychol. Bull.* **51**, 172–5.
STROUD, J. B. & HENDERSON, M. (1943). Rate of reading and learning by reading. *J. Educ. Psychol.* **34**, 193–205.
TINKER, M. A. (1939). Speed versus comprehension in reading as affected by level of difficulty. *J. Educ. Psychol.* **30**, 81–94.
TINKER, M. A. (1940). Dr Robinson on speed versus comprehension in reading: a discussion. *J. Educ. Psychol.* **31**, 559–60.
WOODWORTH, R. S. (1938). *Experimental Psychology.* New York: Holt.

(*Manuscript received* 14 *June* 1957)

EFFECTS OF PRACTICE AND REST ON FLUCTUATIONS IN THE MÜLLER–LYER ILLUSION

By H. J. EYSENCK and P. SLATER

Institute of Psychiatry, University of London

Fifty subjects are given sixty trials each with an apparatus exhibiting the Müller–Lyer illusion. The first forty trials (divided for analysis into four periods of ten) are uninterrupted; a 30 sec. pause for fixation is followed by a fifth period of ten trials; and a rest pause of 30 min. by a final period. On the group as a whole, practice produces scarcely any effect. The means differ very little from trial to trial or from period to period, and the amount of variation about them tends to remain constant. The effect of the illusion varies greatly on different individuals. The variance between the subjects' means for the experiment as a whole constitutes 63% of the total. There are also striking differences in the effects of practice on different subjects; these person:period interactions account for another 14%. The error variance forms almost the whole of the remainder (22%).

To classify the person:period interactions, a principal component analysis was employed. The first latent root, accounting for 59% of the variance due to interaction, describes a course of change which retains the same direction for any particular subject (whether upwards or downwards, towards or away from the point of zero illusion) up to the end of the fifth period but is reversed by the rest pause. The authors do not consider the terminology of satiation or improvement with practice suitable for describing the whole range of such phenomena, but suggest a description in terms of habit reinforcement. However, the phenomena exhibited by subsequent components, which cannot be neglected as statistically non-significant, do not seem so explicable. All such performances constitute patterns of search, and it is noteworthy that none of them accord with the oscillatory pattern suggested by cybernetic theory. Scores on scales of neuroticism and extraversion-introversion failed to correlate significantly with individual differences either in the effect of the illusion over the experiment as a whole or in the progressive changes induced by practice.

I. Introduction

Few experimental variables in the field of learning have been studied more assiduously than the effects of practice and rest upon the development of the particular function under investigation. Experimental work in the field of perception, on the other hand, has not usually been concerned with variables of this type. Possibly due to the phenomenological bias of many workers in this field, massed or spaced repetition of trials has usually only been resorted to in order to increase the reliability of individual observations, and not in order to study the effects of such repetition on the dependent variable.

In previous contributions Eysenck (1955, 1957*a*) has suggested that certain effects in the perceptual field, which Köhler calls *satiation* effects, resemble certain other effects in the field of learning theory, which Hull calls *inhibition* effects. It was further hypothesized that inhibition-satiation effects are more strongly evoked in persons lying on the extraverted end of the extraversion-introversion dimension than those towards the introverted end. This hypothesis has been verified in several studies at an acceptable level of statistical significance (Eysenck, 1957*a*). Satiation phenomena present one of the few instances where the effect of practice upon perception has been studied, and the possibility of finding a link here between learning theory, perception, and personality theory may be worth investigating.

One of the phenomena which may be studied in this connexion is that of the so-called visual illusions. While most interest has, from the beginning, been aroused by the

phenomenon itself and its possible relations to the analytic-synthetic modes of approach of different personality types, several workers have none the less been interested in the course of development taken when the illusion patterns are presented a large number of times. The work of Judd (1902, 1905), Seashore (1908), Lewis (1908) and Brown (1953) on the Müller–Lyer Illusion, that of Cameron & Steele (1905) on the Poggendorf Illusion, and that of Judd & Courten (1905) on the Zoellner Illusion, may be mentioned in this connexion. These writers put forward the general hypothesis that, as Köhler & Fishback (1950*a*, 1950*b*) put it, 'all such illusions represent errors of judgements rather than actual distortion of visual objects. From this point of view, it may seem natural to assume that when subjects deal with the illusion patterns in a great many trials, they gradually learn to avoid these errors, and that the illusions disappear for this reason.' Traditionally, then, although perception is involved in the experiment, the interpretation is in terms of learning, and comparatively little interest would attach to the alleged disappearance of the illusion with practice. But the only type of fluctuation with practice such a simple theory would be adequate to explain, would be a progressive decline in the amount of the illusion from some initially large positive value to a final value approximately zero. This is far from being the only possible origin, direction of change, or termination to be found described in the literature or observed by us.

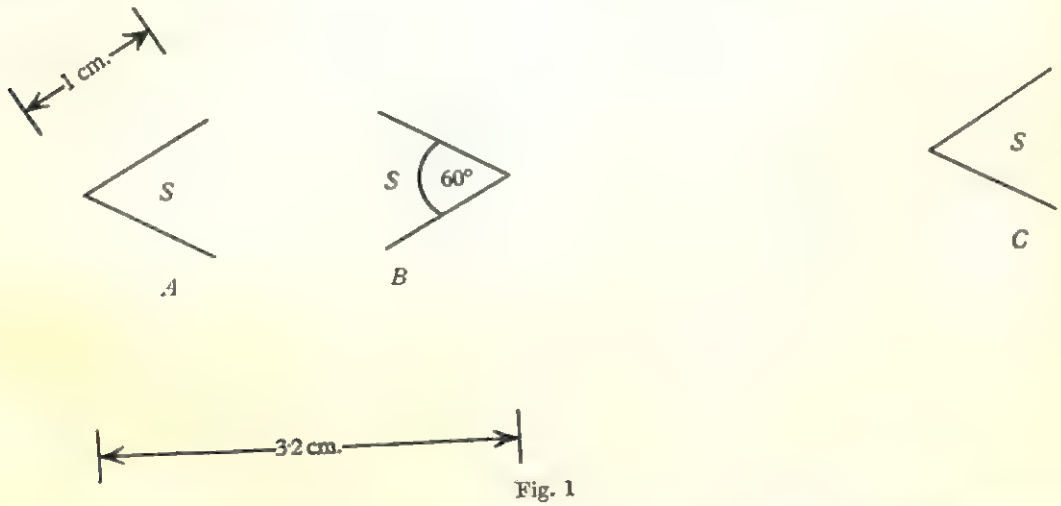


Fig. 1

In a brilliantly argued series of deductions, Köhler & Wallach (1944) and Köhler & Fishback have maintained: (1) that the explanation in terms of learning is inadmissible because subjects do not know where, and to what extent, their judgements are in error, and cannot therefore correct this; and (2) that the phenomenon of the disappearance of the illusion with repeated trials could be accounted for in terms of the general principles of satiation. This explanation, briefly, is that satiation develops more quickly inside the angles formed by the arrowheads, so that, as satiation develops, the apparently shorter distance included within the arrowheads will appear to become extended until it equals, or even exceeds, the distance between the oblique angles. In terms of Fig. 1, satiation develops most strongly at the points marked with an '*S*', pushing, in terms of the subject's visual experience, point *A* to the left, point *B* to the right and point *C* to the left.

So long as the only phenomenon to be explained is a progressive diminution in the effect of the illusion with practice, a choice between 'satiation' and 'learning' as descriptive terms cannot be guided by the evidence, and must rest on the relative strength of the logical arguments in support. But Köhler and his colleagues note a wide variety of other phenomena, and while they adduce such evidence as in conflict with the learning theory, they run into difficulties in attempting to extend their own satiation theory to accommodate it. Particularly with massed practice, and at least in some people, it has been reported that the illusion *increases* rather than *decreases* with time. This is acknowledged by Köhler and Fishback when they say: 'The obstacle which develops when many trials are given without rest periods does not merely delay the process by which the illusion is destroyed. Often its disturbing effect becomes so strong that for a while the direction of the development is actually reversed.' They also quote Seashore's report that one of his subjects showed no change in the size of the illusion even after measurements had been made for 24 days. The modified hypothesis they propose to account for such phenomena is by no means clear-cut and obvious. They point out (footnote on p. 274) that at first sight it might appear that, in such subjects, satiation must be exceptionally weak or that it must fade particularly fast, 'but if we are not mistaken one could also make the opposite assumption. There can be little doubt that, as a consequence of everyday vision, the visual cortex as a whole is in a state of quasi-permanent satiation, and that local satiation as established in special fixation periods is merely added to the level of that persistent condition. . . . Now satiation is a self-limiting process. . . . It therefore seems at least possible that, with some subjects, temporary and local satiation effects are found to be weak because the more persistent satiation of the tissue as a whole is unusually strong.'

We find it very difficult to operate with this modified hypothesis. It seems to mediate alternative expectations without specifying the conditions under which one or another is to be preferred, and thus to be of a kind repugnant to science. Moreover, the properties Köhler attributes to satiation are not at all in agreement with common findings which are that satiation effects disappear relatively rapidly. Long-lasting effects are much more typical of learning phenomena.

Köhler & Fishback's major objection, that when subjects are precluded from comparing their results with the objective situation they cannot learn, implies a restriction to the connotation of the term learning which may not be psychologically tenable. Learning in the sense of habit formation may occur without any such facilities, as long as a *reinforcement* is provided. Köhler himself, in a footnote on p. 278 of his paper with Fishback (footnote 13), has given a clue as to a possible source of reinforcement and our own observations are in agreement with his. This is what he writes: 'When subjects have sufficient time to do so, they will usually try to make sure that their first impression is reliable, and will therefore postpone their final decision. Since, during this period, satiation can change the appearance of the pattern, they may find themselves disturbed by contradictory impressions.' We have frequently found in our experiments that a subject may become discontented with the setting he has chosen while the experimenter is writing it down; if so, a reinforcement is provided for a change in that direction on the next occasion, and a habitual trend of change may thus be built up. This may be towards a decrease in the illusion if the initial effect is positive and the direction taken is downwards; but even so it may continue beyond the point of objective equality and terminate in a negative effect. Or the trend may equally well be in the opposite direction. The relation-

ships of such trends to the point of equality will differ according to their starting points, directions and rates of change; but they may all be instances of habit formation.

However, if an ethical implication is considered inseparable from the term learning, in the sense that, at the end of a learning process, performance must in some way be *better* than at the beginning, its application to the patterns of behaviour we are considering must be withdrawn. In that case, it might be better to employ the term exploration to describe them. The adoption of a plan of search is found to be normal in subjects with a mental age of 13 years or over; and it may therefore be normal for an adult subject, dissatisfied with a particular setting during a series of trials with the Müller-Lyer illusion, to try modifying it systematically rather than at random What we are suggesting then, is that, among such subjects, an orderly trend is observable in exploratory behaviour, not necessarily following the same pattern in all cases; that it may be maintained by some reinforcing stimulus as long as no obstacle or interruption disturbs it; and that there are consequently some similarities between exploratory and learning processes.

Such theories do not enable a particular sequence of performances to be defined as the most probable one for a particular subject on any evidence available prior to the experiment, but lead to some expectations concerning the performances of groups of subjects which can be verified. No general tendency need be expected for the illusion to increase or decrease during a period of massed practice, nor for the variance about the mean effect to increase or diminish; differences in starting points, directions and rates of change, being free to vary for different subjects, will operate as that multitude of small independent causes which tends to stabilize a normal distribution. But the performances of individual subjects may be expected to show systematic changes, demonstrable by dividing a long series of trials into a few periods of several trials each and testing whether the individual's mean performances differ significantly from one period to another. There is no particular reason for expecting that the systematic changes exhibited by different individuals will be related to their scores on measures of extraversion-introversion if the satiation theory is untenable.

On any form of the satiation theory, however much modified, it would seem reasonable to expect some general tendency in a group of subjects towards convergence on the point of objective equality at some time during a long series of trials, more probably nearer the end. If the initial effect of the illusion is large, it should tend to decrease; if small, to remain small or decrease further, possibly to a negative end-point. The exceptional cases should be relatively few.

II. EXPERIMENT

(*a*) *Subjects.* Fifty male subjects between 18 and 25 constituted the experimental group; nearly all were University students and all were experimentally naïve and had no knowledge of the aim of the investigation.

(*b*) *Apparatus.* An illusion pattern was housed in a wooden box, 45 in. long, 4 in. high and 6 in wide. A special face-mask made of rubber enabled the subject to press his face against the open end of the box to the virtual exclusion of all extraneous light, thus reducing external clues. The inside of the box was painted matt black and the pattern was presented to the subject at the other end. It was illuminated from behind by means of a 25 W. lamp. The pattern was drawn on a double sheet of Perspex and was presented with the subjectively shorter distance on the left; this distance was kept constant. The subjectively longer distance could be varied by pulling out the part of the Perspex on which it was drawn and pushing it back again. The dimensions of the illusion, which was presented simply in terms of the angles and without a central line to connect them, were as shown in Fig. 1.

(*c*) *Technique of measurement.* The experimenter pulled the adjustable part of the apparatus out to a point where preliminary experiments had shown there would be no chance at all of any subject regarding the distance as equal. He then pushed it very slowly back until the subject said 'halt'. His instructions were to call out when the distance between the points of angles *A* and *B* appeared equal to the distance between the points of angles *B* and *C*. The experimenter recorded the amount of error shown by the setting (difference of constant minus adjustable distance in mm.) and removed the adjustable part again to an extreme position somewhat different from that previously used. The subject kept his eyes open throughout and fixated the angle at *B* in Fig. 1. Then the experimenter repeated the experiment immediately, leaving as little time as possible between successive judgements. Forty judgements were made in a row numbered below 1–40. Then the subject was instructed to fixate the illusion pattern by keeping his eye fixed on the point of the angle marked *B* in Fig. 1; this fixation had to be held for 30 sec., after which another ten readings were taken—numbered 41–50. After the fiftieth reading, a 30 min. rest pause was introduced; then another ten readings were taken, numbered 51–60.

(*d*) *Instructions.* The subjects were told throughout to fixate the point of the angle marked *B*. They were not told of the nature of the illusion or the fact that an illusion was being presented to them at all. They were simply asked to estimate the distances between three points at the tips of the angles and to say halt whenever they considered the distances to be equal. They were shown the mechanism before starting and none made any errors during the run of the experiments. All subjects filled in an extraversion-introversion and neuroticism questionnaire, specially designed by one of us (H.J.E.) and described briefly in another publication (Eysenck, 1957*b*). It was considered possible that the 'analytic' and 'synthetic' types of subjects mentioned in the literature might be found to differ from each other in one or both of the personality dimensions, but no scale relating directly to this trait was used.

III. Description and analysis of the results

The group mean for the entire experiment, i.e. the mean of all 3000 measurements, is 1·551 mm., and the sum of squares of deviations from it is 66,106·09 (cf. Table 2). The average effect of the illusion is thus not great compared with the amount of variation about it. Readings range from +16 to −14 mm. and 39·8% are negative. Their standard

Table 1. *Results illustrating periodic fluctuations in the effects of the illusion*

During period	All subjects		Means of subjects		
	Mean	S.D.	S1	S2	S3
1	1·14	4·88	9·8	2·9	−8·4
2	1·28	4·78	9·2	4·0	−8·6
3	1·75	4·44	4·2	2·7	−6·8
4	2·13	4·61	1·7	4·5	−3·9
5	1·56	4·90	0·2	3·4	−1·4
6	1·43	4·48	2·0	3·1	−5·6
All periods	1·55	4·69	4·52	3·43	−5·78

Table 2. *Analysis of the variance of the measurements*

Source	Sum of squares	Degrees of freedom	Mean square variance
Total variance	66,106·09	2999	850·40
Variance between—Persons	41,669·53	49	
Periods	315·57	5	63·11
Person : period interactions			
First latent root	5620·50	49	114·70
Second latent root	1991·07	49	40·63
Third latent root	960·14	49	19·59
Fourth latent root	616·17	49	12·57
Fifth latent root	302·99	49	6·18
Total	9,488·99	245	38·73
Variance between trials within period	248·80	54	4·61
Person : trial interactions within period	14,383·20	2646	5·44

deviation is 4·692 mm. The form of the distribution approximates to the normal, though it does not conform exactly. The value of β_1, −0·1933, indicates slight negative skewing (mean below mode), the maximum frequencies occurring at 2 and 3 mm. The value of β_2, 2·561, shows that the form is also slightly platykurtic.

The individual means for the entire experiment, obtained by summing the sixty measurements for each subject, have a very wide range, from +8·90 to −5·78 mm.; fourteen are negative. Their sum of squares, 41,669·53, accounts for 63% of the total variance. Subjects thus differ greatly in their susceptibility to the illusion; and such individual differences are the most important observations for any adequate theory of the Müller–Lyer illusion to explain. Their relationship to our personality measurements is very slight, if any: they correlate 0·1860 with the extraversion scale and −0·1151 with the scale for neuroticism. Neither correlation is significant for fifty cases.

The literature on the relation between susceptibility to the illusion and personality (Schiller, 1942) suggests that it is the synthetic type of person who is susceptible to the illusion as opposed to the analytic type. Personality descriptions given by continental authors suggest that the synthetic type is more extraverted than the analytic type, as well as less neurotic. The direction of the two correlations is in line with this hypothesis, but only the former appears worthy of further investigation, being almost significant on the one-tail test. Accordingly, our subjects were divided into three groups, made up of the ten most extraverted subjects ($E+$), the ten least extraverted subjects ($E-$), and the intermediate thirty subjects ($E0$). The mean errors for these three groups were 3·525 for the $E+$ group, 0·820 for the $E-$ group, and 1·142 for the $E0$ group. This order of the means is in line with the hypothesis. Analysis of variance, however, gives a P value somewhat higher than 0·05. The data may be suggestive, particularly in view of the fact that the order of the means would agree with prediction by chance only in one case out of six, a fact which is not taken into account by the analysis of variance technique. Further work along these lines might be promising.

To distinguish between the irregularities and the systematic changes in the measurements obtained during the course of the experiment the sixty trials have been divided into six periods of ten. Periods 1 to 4 succeed uninterruptedly; then comes the 30 sec. pause for fixation; then period 5; then the rest pause; then period 6. The variation between periods registers the relatively long-term changes, and the variation between trials within periods the momentary irregularities in individual performances. A comparison of the mean-square variances (M.S.V.) from these two sources excludes some classes of hypothesis altogether, though it leaves the choice between others open.

If the two M.S.V.'s do not differ significantly, no evidence of any systematic component is disclosed. If the M.S.V. between periods is significantly greater than the M.S.V. between trials within period, the evidence is consistent with some progressive change of the kind discussed in the introduction. If the latter exceeds the former, oscillation of some kind is suggested: this might be consistent with an attempt to arrive at the true length by cybernetic methods, overshooting it alternately in opposite directions.

In calculating the total variance (sum of squares) for this comparison, the sum assigned in Table 2 to person:period interactions should be included with the variance between periods; and the person:trial interactions within period are to be included with the variance between trials within period. The two M.S.V.'s are 39·22 with 250 D.F. and 5·42 with 2700 D.F., respectively. As the former is very much the greater, progressive change of some kind is indicated; neither of the alternative kinds of hypothesis is acceptable.

If a single direction or course (not necessarily common to all the subjects or rectilinear over the entire experiment, but sufficiently common and consistent to be statistically

significant) is to be found in the progressive changes in the measurements obtained from different subjects during successive periods, it will be shown in the group means per period. But if the directions taken by different subjects vary independently, they will tend to obliterate one another, and the variation in the group means will tend to be small and erratic. Thus consistent changes will augment the variance between periods and independent changes will augment the person : period interactions.

The observed effects are illustrated in Table 1, where the period means for the first three subjects are tabulated alongside the means for the entire group. S1, as can be seen, is markedly affected by the illusion at the beginning of the experiment; as it proceeds, the illusion appears to wear off and during period 5 its effect is very slight; after the rest pause it appears to have returned, though not in its original strength. A graph of this performance would resemble Köhler's Fig. 3 (1950*a*) very closely. S3's performance is almost the exact reverse: he starts with a substantial underestimate of the distance the illusion supposedly lengthens; his estimates increase as the experiment proceeds, approaching the true value most closely again at period 5; and again, after the rest pause, there is a relapse. But no such course is followed by S2; in his case the illusion appears to be comparatively stable. The group means, which are the net resultant of such divergent courses, also show little fluctuation.

That the courses followed by different subjects vary from one another and do not cumulatively define any single predominant direction of change is shown by the fact that the M.S.V. between periods (due to differences in the group means per period) is not significantly greater than the M.S.V. of the person : period interactions. This disposes of the simpler forms of the learning and the satiation theories discussed in the introduction, together with any other theories which postulate that a single course of change will be found to predominate during a period of massed practice. Even it it succeeded in accounting exactly for all the variations in mean measurements from trial to trial as well as from period to period, any such theory would leave over 99% of the total variance unaccounted for: the variance between the sixty trials is only 0·85% of the total.

To discover any common characteristics among the courses followed by different subjects it is necessary to analyse the person : period interactions. Next to the differences between individual experimental means, they are the largest source of variance outside the experimental error, comprising 14·4% of the total. Detailed tabulations were made. For example, S1's mean for the entire experiment is +4·52 (cf. Table 1); it exceeds the group mean by 2·97 mm. If his course had conformed with the group's, his means per period would have been 4·11, 4·25, 4·72, 5·10, 4·53 and 4·40. The deviations of his observed means from these expectations, viz: +5·69, +4·95... −2·40, measure his period interactions. Tabulating them for each subject provides a matrix of 6×50 measurements with, however, only 5×49 D.F. between them. Significant interactions were not found to be confined to a few of the subjects. They lie beyond the 0·01 probability limit in thirty-three cases and between the 0·05 and 0·01 limits in three more; it is the absence of any interaction, as in the case of S2, which is unusual.

Their variances and covariances per period are given in Table 3. Looking along the leading diagonal it is noticeable that the largest deviations occur towards the beginning and end of the experiment; the smallest variance is at period 3. Deviations in the first three periods tend to correlate positively with one another and negatively with those in subsequent periods. There is evidently a general tendency for the subjects to give responses

at the start of the experiment which lie at a considerable distance away from what subsequently prove to be their mean points for the experiment as a whole; to converge on these points as the experiment proceeds, coming nearest usually in period 3; and continuing roughly in the same direction to move on past them away again. Subjects 1 and 3 (cf. Table 1) follow courses which conform approximately with this description.

Table 3. *Matrix of the total variances and covariances* of the person : period interactions*

3309·5	762·5	−262·7	−1574·1	−1737·2	−498·4
—	1092·5	109·5	−456·8	−910·6	−597·6
—	—	617·5	291·5	−306·5	−448·8
—	—	—	1359·3	331·6	47·8
—	—	—	—	2118·8	503·9
—	—	—	—	—	993·3

* Sums of squares and products, with 49 D.F.

It might be supposed, if this is the general tendency, that there will be some point during the experiment when the measurements obtained lie closer, if not to the true value, at least to the experimental mean for the group. But this is not the case. As the S.D.'s noted in Table 1 show, the amount of variation about the period mean for this group is approximately the same at all stages during the experiment. This is because the convergence which tends to occur over the first thirty trials is of the individual's responses towards their own experimental means, which vary widely about the mean of the group; nor is period 3 always the point at which convergence is closest; and indeed persistence in the same direction does not remain uninterrupted in every case.

To reduce vagueness in discussing what general tendency or tendencies are to be found among these interactions a technique is needed for abstracting common characteristics and measuring their importance. An adaptation of principal component analysis is submitted by one of us (P.S.) for this purpose. It leads to the latent roots of the components included in Table 2 and the weights proportionate to factor loadings given in Table 4.

Table 4. *Latent vectors of the person : period interactions**

CI	CII	CIII	CIV	CV
0·731	−0·309	0·107	−0·371	−0·231
0·279	0·311	−0·421	0·686	−0·107
−0·007	0·429	−0·102	−0·387	0·700
−0·341	0·469	0·413	−0·150	−0·552
−0·492	−0·463	−0·560	−0·201	−0·151
−0·171	−0·436	0·563	0·424	0·344

* Scaled appropriately for calculating component measures from preceding residuals. To obtain loadings for extracting components from the matrix of total variances and covariances (Table 3) multiply each vector set by the square root of its latent root, given in Table 2.

When multiplied into S1's interaction measurements, the weight for the first component gives his C1 index, thus:

$$5{\cdot}69 \times 0{\cdot}73 + \ldots + (-2{\cdot}40) \times (-0{\cdot}17) = 9{\cdot}244$$

and the C1 index, multiplied by the weights, gives the modified expectation of his performance during the successive periods illustrated in Fig. 2. The fit is evidently improved substantially in this case.

The C1 index for the group as a whole varies about 0 with a S.D. of 3·387, i.e. $(5620{\cdot}50 \div 490)^{\frac{1}{2}}$. Particular values designate individual specimens of the same general type of

performance, analogous to members of a family of curves designated by a particular choice of constants. A positive index describes a course dropping, like S1's, to a trough at period 5, but the point of the trough may be located anywhere above or below the base line of equality (no illusion). Conversely a negative index describes a course rising, like S3's, to a crest at period 5; and this too may be located anywhere above or below the base line. The position of trough or crest relatively to the base line depends, for any given CI index, on the subject's experimental mean.

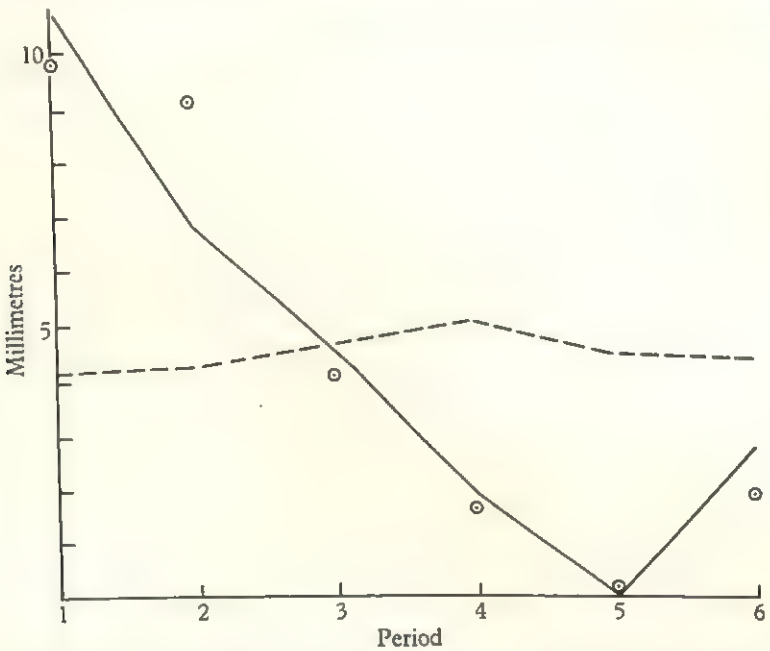


Fig. 2. Period means for S1. ⊙, Observed; – – –, expected from group means; ——, expected taking CI into account.

While the terminology of satiation may apply approximately to the course followed by S1, trying to extend it to S3 may be felt to strain it. The terminology of learning, in the sense of improvement with practice, might apply equally well to both subjects' performances; but it too would be strained if extended to cases where a positive experimental mean occurs concomitantly with a negative CI index, or a negative mean with a positive CI index, i.e. where the course towards the crest or trough takes the subject *away* from the base line. Such occurrences could be readily exemplified from our data, and the statistical analysis shows that they can be regularly expected and all referred to a common form. However, in terms of the modified learning theory outlined in the introduction, that is to say as phenomena of simple habit reinforcement, all courses of the CI type are equally explicable, and the reversal of direction which occurs after the 30 min. rest pause is also intelligible. If a psychological description is to be attempted, these terms would seem preferable.

Expectations derived from CI naturally do not coincide with the observed measurements at every point in every case; the deviations from expectation give rise to a matrix of residual variances and covariances from which a second principal component, CII, can be extracted. The process concludes only after the extraction of five components. Just

as the CI index modifies the expectations based only on the experimental and group means, the CII index modifies again the modifications introduced by CI until, finally, with CV the whole variance of the interactions is taken up, and an exact fit is obtained to every subject's observed mean performance per period. The significance of the improvement in fit provided by each modification is tested in the analysis of variance (Table 2). CV is the only component that falls short of significance at the 0·01 probability level.

There thus appear to be types of course which deviate significantly from CI and are not readily ascribable to habit reinforcement. In particular, the 30 sec. fixation period produces startling changes in courses of both the CII and CIII type which are not readily explicable in terms of any theory we have discussed. In accordance with satiation theory the prevalent tendency (downwards towards the base line) should be accelerated, while in terms of simple habit reinforcement no large change is to be expected: but it is evident from CII and CIII that complete reversals are quite common. While all the components could be described as classifications of common plans of search, it is not immediately evident that such a terminology will be helpful for purposes of inference.

One possible hypothesis to account for the reversal in the direction of change during the fixation period might be constructed with reference to reactive inhibition of those cortical elements active in the attention and judgement processes which are supposed to characterize the 'analytic' type of person. Such inhibition might be supposed to accumulate during the forty massed practice trials preceding the 30 sec. inspection period, and, apparently, to dissipate during this period. Experimental evidence is available to show the existence of reactive inhibition in relation to such cortical processes and a direction test of this hypothesis would not seem to be impossible (Eysenck, 1957*a*).

There is no evidence of any significant association between any of the components and scores on the scales for neuroticism or extraversion-introversion. The correlations are

	Scale	
Component	*N*	*E*
I	0·0804	0·0376
II	−0·1842	0·1711
III	0·2138	0·0606
IV	0·1905	0·0674
V	0·0871	−0·0864
All components multiple *R*	0·3605	0·2153

The 0·05 probability limit for a zero correlation is ±0·2829 in a sample of fifty cases and 0·5255 for a multiple correlation based on five independent variables. There may well be some other personality trait or traits with which these components are associated, e.g. the analytic/synthetic difference. But of this we have no evidence. We cannot at the moment say whether the components described have any significance for personality theory at all.

REFERENCES

BROWN, K. T. (1953). Methodology for studying figural after-effects and practice effects in the Müller–Lyer illusion. *Amer. J. Psychol.* **66**, 629–34.

CAMERON, E. H. & STEELE, W. N. (1905). The Poggendorf illusion. *Psychol. Rev.* (*Mon. Suppl.*), **7**, 83–111.

EYSENCK, H. J. (1955). Cortical inhibition, figural after-effect, and the theory of personality. *J. Abn. Soc. Psychol.* **51**, 94–106.

EYSENCK, H. J. (1957*a*). *The Dynamics of Anxiety and Hysteria.* London: Routledge and Kegan Paul.

EYSENCK, H. J. (1957*b*). The questionnaire measurement of neuroticism and extraversion. *Riv. Psicol.*, **50**, 113–140.

JUDD, C. H. (1902). Practice and its effects on the perception of illusion. *Psychol. Rev.* **9**, 27–39.

JUDD, C. H. (1905). The Müller–Lyer illusion. *Psychol. Rev.* (*Mon. Suppl.*), **7**, 55–81.

JUDD, C. H. & COURTEN, H. C. (1905). The Zöllner illusion. *Psychol. Rev.* (*Mon. Suppl.*), **7**, 12–39.

KÖHLER, W. & FISHBACK, J. (1950*a*). The destruction of the Müller–Lyer illusion in repeated trials. I. An examination of two theories. *J. exp. Psychol.* **40**, 267–81.

KÖHLER, W. & FISHBACK, J. (1950*b*). The destruction of the Müller–Lyer illusion in repeated trials. II. Satiation patterns and memory traces. *J. Exp. Psychol.* **40**, 398–410.

KÖHLER, W. & WALLACH, H. (1944). Figural after-effects. *Proc. Amer. Phil. Soc.* **88**, 265–357.

LEWIS, E. O. (1908). The effect of practice on the perception of the Müller–Lyer illusion. *Brit. J. Psychol.* **2**, 294–306.

SCHILLER, L. (1942). Ganzheitliche Auffassung und Persönlichkeitstypus. *R. Psychol.* **153**, 43–80.

SEASHORE, C. E. and others (1908). The effects of practice on normal illusions. *Psychol. Rev.* (*Mon. Suppl.*), **9**, 103–48.

(*Manuscript received* 8 *July* 1957)

PUBLICATIONS RECENTLY RECEIVED

Principles of Perception. By S. Howard Bartley. New York: Harper. London: Hamish Hamilton. 1958. Pp. xii+482. 52*s*.

It is perhaps unfortunate that Dr Bartley chose, as the title of his book, *Principles of Perception*, since although he does enunciate certain useful principles to be observed in studying and hypothesizing about perception, the greater part of the book consists of an introduction to the subject describing some of the experimental investigations which have been carried out. Inevitably only a selection of these could be covered, since all the sensory modes are included; and there is a bias towards experimental work on the physiological psychology of perception, although little purely physiological theory is discussed. Thus, colour vision and hearing are left somewhat in the air. However, obviously the book would have been excessively long had their theory been included.

A more serious fault is the rather scrappy treatment of certain of the psychological phenomena of perception, which seems to be due in part to the omission of almost any reference to the work of other than American psychologists. The psychology of form perception and the Gestalt studies are hardly mentioned, and the work of Piaget and Michotte is omitted altogether. It is ludicrous to devote a chapter to the 'development of perception' with no reference to Piaget. However, there is an excellent critical discussion of the role of social and motivational factors in perception, and particularly of the number and vagueness of the hypothetical constructs, such as 'subception' and 'perceptual defence', which have been introduced in connexion with these studies.

But this book provides a good introduction to the study of perception, provided that it is not regarded as comprehensive. It will be the more valuable if some of the principles stated by the author are understood and applied in studying and performing experimental work on perception. In the first place, he stresses the distinction which should be made between the qualities of the physical stimulus and the nomenclature applied to them, and the qualities of the perceptual experience with its appropriate terminology. Not only must the student have this distinction clearly in mind; but he must also differentiate between the essential discriminatory characteristics of the immediate perception, and those processes appertaining to the further elaborations of judgement and valuation, which in the past have often been confused with them. The trouble, of course, comes in knowing where to place the processes of inference and identification which generally form an integral part of perception; and the author is not always very clear in his discussion of these. He does, however, make the important point that certain characteristics are merely descriptions of the percepts, and cannot be regarded as causal factors determining them. This raises the issue of the functions of 'cues', for instance, 'cues' to the perception of distance. Unfortunately, this issue is not adequately clarified, and no discussion is included of observations made on the effects on spatial perception of prolonged inversion or reversal of the visual field. Instead, the author wanders off into a lengthy discussion of physiological optics which seems to be of little interest to the student of psychology.

Some of the main theories of perception discussed by F. Allport are briefly stated at the beginning of the book and recapitulated in the final chapter. The author points out that though several of these can be applied to certain particular perceptual phenomena, none accounts satisfactorily for all such phenomena. But this seems inevitable and even desirable in the present state of our knowledge. It is far better for the beginning student to be confronted with the gaps and inconsistencies between rival theories than to suppose that there can as yet be any definitive and all-inclusive theory to which all the mass of incoming experimental data must conform.

M. D. VERNON

Contemporary Approaches to Cognition. By J. S. Bruner *et al.* Harvard University Press. 1957. London: Oxford University Press. 1958. Pp. vi+210. 32*s*.

This record of Colorado University's 1955 symposium on cognition contains a short introduction by Muenzinger and six chapters, each by a different author and all but the last followed by a critique from one of the other participants. The first chapter, by the late Egon Brunswick, maps out the place of cognition in that polysyllabic conceptualization of organism-environment interaction made familiar by Brunswick's previous writings. Then follows what is, for the reviewer, perhaps the most stimulating

contribution; Bruner's 'Going beyond the information given'. Taking his cue from Bartlett's schemata and extending his own work on categorizing, Bruner outlines a scheme which, whatever its limitations, lends unity to a wide array of phenomena. His integrating concept is 'the coding system', i.e. that way in which people assimilate information to organized bodies of past experience and, in terms of this assimilation, process and go beyond what is given. If a person is given the numbers 8 and 4 and asked 'What is the answer?', this answer will depend on whether the thinker superimposes on the data the templet of multiplication, division, addition, subtration, or continuation of a series: he codes the data so as to extrapolate from them in one way rather than another. This example is triviala nd also caricatures the extent to which most problem situations are ambiguous but, in the hands of Bruner, the notion which it illustrates is elegantly applied to much of perceiving, learning and thinking: and in the later clinically oriented chapter by Rapaport on 'Cognitive structures', this same notion may be seen to extend to those quasi-permanent cognitive 'styles' and value-systems of adult personality. The longest chapter is Osgood's bold and sparkling 'Behavioristic analysis of perception and language' which ranges all the way from sensory mechanisms to meaning by a breath-taking *tour de force*. Festinger contributes a cogent discussion of what he calls 'cognitive dissonance' (his book on this topic has been published in the interim), and Fritz Heider rounds off with a chapter on 'Trends in cognitive theory'.

This book is unlikely to be read continuously for many years to come. Nor does it set forth all contemporary approaches to cognition: two notable omissions are Piaget's developmental work and the case-history study of the circumstances leading up to scientific invention and discovery. The book's importance lies in its timeliness. It brings together investigators of varying backgrounds for free and speculative discussion on cognitive processes at a time when psychologists are evincing a growing interest in these processes without, as yet, having split into opposed theoretical camps; and its proceedings may awake in many readers some appreciation of cognitive problems and perhaps influence some psychologists to tackle these problems in a many-sided and non-partisan way. If any facet of human activity is, more than any other, the proper subject matter of psychology, then cognitive activity is surely it. Yet following their early concern with 'higher mental processes', psychologists have, with a few noteworthy exceptions, shown little interest in this facet: the study of reasoning, thinking and imagining has gone through a 'latent period' during most of the first half of the present century. It might, with some justification, be claimed that this latency reflects lack of adequate methodological and conceptual tools and that these tools have to be furnished, somewhat indirectly, from explorations of: traditional sensory and learning laboratory experiments with animals and humans; statistical control of and inference about mental test performances; the properties of mechanical and electronic systems; the often bizarre interpersonal to and fro of psychotherapy; and accumulated naturalistic observation. It might further be claimed that, armed with these tools, psychologists will return to 'higher mental processes' with an attack which is more varied, subtle and fruitful than before. The most forceful overall impression made by this book is that this renewed and more effective attack has begun in earnest.

I. M. L. HUNTER

A Theory of Cognitive Dissonance. By LEON FESTINGER. Illinois: Row, Peterson and Co. 1957. Pp. ix + 291. $ 5.25.

Where there is inconsistency between the various attitudes and beliefs of an individual there will also be what Festinger calls 'psychological discomfort'. It is this inconsistency resulting in psychological discomfort which is discussed in the book under the label of 'dissonance'. Likewise, the word 'consistency' is replaced by the word 'consonance'. From two basic hypotheses—that where there is dissonance a person will be motivated to try to reduce it, and that when such dissonance is present a person will actively avoid situations and information which would increase the dissonance—Festinger proceeds to develop his theory of cognitive dissonance. Thus the basic data of this theory is common property to all men and some might say furthermore that common sense will take us far in understanding the way in which such dissonance arises and is changed. It is Festinger's conviction that it is possible to make this notion of cognitive dissonance scientifically respectable and moreover that, by avoiding loose usage of the concept of dissonance, it is possible to construct a theory which can make specific predictions which can then be experimentally verified or otherwise. The book is equally divided between theory and data in support of the theory. The data examined in the light of the theory are of two classes: on the one hand, social survey data and, on the other hand, data derived from controlled laboratory experiments with small groups. This reviewer finds the evidence from the laboratory experiments much more unequivocal in support of the theory than the social survey data. Moreover, one is impressed by the ingenuity shown

by Festinger and his collaborators in their attempts at the precise measurement of somewhat nebulous entities. It is refreshing to find an author so ready to claim no more than Festinger does when, in his closing sentence, he writes 'If treated as precisely and specifically as possible, I believe the theory of cognitive dissonance will prove a useful explanatory and predictive device'. This book is to be recommended primarily to those interested in social psychology and sociology rather than in the field of cognitive psychology.

M. A. JEEVES

Introduction to Method in Psychology. By W. M. O'Neil. Melbourne University Press. London: Cambridge University Press. 1957. Pp. ix+155. 25*s*.

This is at once a good and exasperating piece of work. It is good because, in a subject plagued by prattling textbooks, it treats the student as a grown-up, and demands of him mental efforts which are justified by the difficulty of the subject matter and necessary to real understanding. It is exasperating because, though it discusses method and system, it seems itself to have been put together in a rather unsystematic way. A concrete example, the 'moon illusion', is described at the beginning as though it were basic for the rest of the book. It is analysed at length in the sections which follow immediately, and returned to here and there during the later stages. But in the meantime other examples have been chosen, other kinds of problem discussed, and the impression left on the reader at the end is not so much of a single text as of a series of penetrating papers dealing with various aspects of scientific method in general. The range is very wide. Particular techniques are dealt with as well as the logic of experiment, and the syllogism finds a place along with analysis of variance and problems of sampling. The level of difficulty throughout the book is uneven. Sometimes the student has to work really hard, later he may have to deal with quite elementary concepts, and then suddenly he is stretched once more. To get most out of it, and there is a lot to get, he should treat the text as something to be referred to in connexion with particular kinds of problem rather than as a connected treatise to be studied on its own. In this way Prof. O'Neil's thinking can be brought to bear on topics as they arise. There is certainly no alternative study of scientific method in the psychological field which has packed so much into such modest space.

JAMES DREVER

Experimental Design in Psychology and the Medical Sciences. By A. E. Maxwell. London: Methuen. 1958. Pp. 147. 17*s*. 6*d*.

It is seldom that one encounters a book that so clearly calls for no review other than an indication of contents and a strong recommendation. Dr Maxwell's concern is with statistical designs, more particularly those likely to be of use to the clinical psychologist in the testing of hypotheses. All are based ultimately on analysis of variance: correlational techniques and the concept of regression are introduced relatively late in the book, at the point at which they become relevant in relation to analysis of covariance. Two designs are included which, as the author points out, have not hitherto been treated in any textbook; these are designs described originally by D. R. Cox in *Psychometrika* (1951, 1955), the more obviously important of which provides a means of handling differences due to irreversible conditions, e.g. effects of a surgical operation.

To claim that the book reads like a novel is perhaps to exaggerate, but the reviewer was unable to put it down while reading at least Chapter I, with its entertaining 'Anecdote concerning Clocks'; the progressive revelation, too, of how the efficiency of a design can be increased with decrease in error variance is not without its element of suspense. And certainly the simplicity and unity of the principles on which the system is based measure up to the criterion of mathematical elegance.

In short, this is both an important and a pleasing book.

B. SEMEONOFF

Elements of Psychology. By David Krech and Richard S. Crutchfield. New York: Knopf. 1958. Pp. xxi+700+xxxvi. $ 8.75.

This is a large new introductory text written mainly from a phenomenological point of view but covering the whole of psychology. The book is divided into four parts headed *Perception* (184 pp.), *Motivation and Emotion* (158 pp.), *Adaptive Behaviour* (148 pp.) and *The Individual* (188 pp.). The parts comprise twenty-five chapters, three of them dealing with physiology and one with psychological measurement.

The contents of the book should please those who are tired of neobehaviourism, learning theory, operationism, and the like. Wherever possible the primary treatment of each topic is phenomenological and, by preference, the authors relate their terms directly to conscious experience rather than to behaviour. As expected, this approach is most successful in dealing with perception and problem solving. The chapter on emotions consists mostly of a classification and the description of conscious states of grief, shame, delight, loneliness, etc. Motivation is dealt with mainly in terms of perceived needs and desires although it is acknowledged that there may be discrepancies between the motive a subject reports and what an external observer would judge his motive to be. In discussing learning, the conditioned response is treated as anticipatory, adaptive, and characterized by 'relative lack of awareness and control'; rote learning is given a rather distinctive perceptually oriented treatment; trial and error learning is taken to be a process of cognitive discovery; human skills are regarded as improving with practice due to the acquisition of 'kinesthetic knowledge'. A distinction is drawn between 'Habitual Behavior', consisting of well-practised sequences of responses going off with a minimum of awareness and 'Coping Behavior', concerned with discovering and attaining goals, and involving attention. Possibly, as the result of concentrating on mental events, the chapters on physiology do not form an integral part of the text. Significantly, these chapters occur at the end of each section. The forty pages on the physiological basis of perception do not seem to be needed at all for the discussion of perception that precedes them. A chapter on measurement effectively examines some problems of scaling and sampling and presents elementary statistics, although it is a little unbusiness-like in that it gives a full discussion of the standard error at the end of twenty-eight pages, but does not even mention significance levels or critical ratios. The remaining chapters, concerned with topics such as growth, intelligence, and personality are all interesting and competent.

One of the best features of the presentation is the use of 'boxes'—material set in smaller type and segregated from the rest of the text. There are 169 such 'boxes' mostly containing reports of experiments and aiming to provide enough detail for the reader to be able to evaluate some of the generalizations that appear in the main body of the book. Line illustrations of a serious and helpful type have been used freely. Informal summaries are frequent and each chapter ends with a full glossary that might be valuable for self-testing. The style of writing is more modest than that of the earlier book by the same authors although words are still used lavishly.

As an introductory text, the book is attractive for a number of reasons. It acquaints the reader with a wide variety of current and 'classical' research; its presentation is intrinsically interesting and serious in tone; its approach is probably more meaningful and acceptable to the beginner than neo-behaviourism; it contains a good deal of sound commonsense. However, as basic introductory reading it is expensive, and even those who favour phenomenology may find that it is in some ways too definitely the product of its authors to be fully satisfactory. Some common terms are given unusal definitions, for example, needs are defined as *experiences* of deficiency; and a number of generalizations, for example, on the nature of discrimination learning, are stated without even a warning that they are controversial.

On the whole, it is probably safe to suppose that those who liked the earlier text on social psychology will like this new book also.

R. L. REID

The Unconscious: a Conceptual Analysis. By A. C. MacIntyre. London: Routledge and Kegan Paul. New York: Humanities Press. 1958. Pp. ix + 100. 11*s.* 6*d.*

Oppressed as we are by long and unreadable books it seems a little churlish to criticize this short and stimulating essay for its sins of omission. But useful as it is, it is not quite useful enough. In a prolegomena and four brief chapters, Mr MacIntyre underlines the extent to which Freud was apt to treat 'the unconscious' as an entity. He then argues that, stripped of substantial peculiarities and re-expressed in adverbial or adjectival form, concepts such as 'repression' and 'unconscious functioning' are so interrelated that Freud's contribution to theoretical psychology should be seen as brilliant and insightful redescription rather than explanation. In his last chapter, Mr MacIntyre seeks to reinforce this view by maintaining that Freudian theory is, in certain respects, incapable of falsification and therefore cannot be granted the status of an explanatory hypothesis.

While all this is stimulating, is it in fact true and if true does it amount to very much? Psychologists have known for long that Freud inclined too much to think in spatial metaphor. This Mr MacIntyre relates neatly to Freud's early neurology and to traces of Cartesian and Lockian views of the 'mind' and 'ideas'. But the nineteenth-century background on which Freud avowedly drew was richer and more

interesting than might be supposed from Mr MacIntyre's pages. Having gone so far he ought possibly to have described the background more fully. The same type of comment applies to the fairly obvious point that the determinism Freud derived from nineteenth-century physics simply doesn't fit with the enlightened self-determination he sought as a result of therapy. But there is more than a simple logical fallacy involved in Freud's position.

More fundamental is the status of Freudian theory as an explanatory hypothesis. Mr MacIntyre may be right in thinking that what we have is brilliant description entangled with views about the interrelations of childhood and adult experience that can be established only by still more empirical inquiry. But, to this reviewer at any rate, the nature of Mr MacIntyre's distinction between describing and explaining is not clear. Hence, it is not clear that to say that Freud 'described' is quite so fundamental a comment as might be supposed. Neither is it clear why, in his last chapter, Mr MacIntyre is so easily frightened by the familiar bogeys of suggestion, imaginary recollection, and ambivalent love and hate, into thinking that a Freudian interpretation could never be falsified with reasonable probablity in a given case. One suspects that 'reasonable probability' as the upshot of inquiry is something which Mr MacIntyre would find unsatisfying.

But to make these comments is perhaps to ask Mr MacIntyre to have written a different book, instead of being grateful, as one must be, to a philosopher for stating some of his perplexities so well and so shortly that they must present a challenge to needlessly muddled thought.

JOAN WYNN REEVES

Personality and Temperament. By SOLOMON DIAMOND. New York: Harper Bros. London Hamish Hamilton. 1958. Pp. vii+463. 48*s*.

In the earlier part of this book, the author is concerned to show that what he calls 'fearfulness', 'aggressivity', 'affiliativeness' and 'inhibitory control' are fundamental dimensions of temperament in both animals and humans. Several chapters are devoted to animals, which is probably no bad thing. There appears to be some confusion over motivation, when it is said that 'drive' depends upon physical tension arising from a 'basic life process', while 'motive' is the result of learning—'thus the "drive" is primary and the "motive" is secondary, while "motivation" embraces both'. A good deal of attention is given to Sheldon's typology, which the author thinks has been unduly neglected. Relatively little attention is given to evaluating techniques for assessing personality, and this being so, perhaps the MMPI comes in for a disproportionate share. Nor is much attention given to the moral aspects of personality, except in a rather superficial way. Of the later chapters, that dealing with the formation of the self-concept curiously lacks any mention of George Mead and any reference to role-playing. Freud, Dollard and Miller, the neo-Freudians, Lewin, and Jung are discussed, but not in a very well-integrated way. In the final chapter on anxiety and integration, several theories are mentioned, but there is no real attempt to assess or integrate these, though the author seems to have a preference for Freud. This relative lack of critical assessment, indeed, is to be found in other chapters as well, as for example in the discussion of the effects of 'emotional deprivation' in infancy.

There is a marked paucity of references to work other than American, and some expressions, such as a reference to *The Bridge of San Luis Rey* as 'Wilder's beautiful novelette' will irritate many British (and no doubt other) readers. On the whole the book, especially the later part, seems to suffer from lack of unity and continuity, and almost creates the impression of consisting of a series of more or less independent contributions, of somewhat uneven merit.

D. GRAHAM

Homosexuality: its Nature, Causation and Treatment. By CLIFFORD ALLEN. London: Staples. 1958. Pp. 143. 15*s*.

Homosexuality. Ed. by CHARLES BERG and A. M. KRICH. London: Allen and Unwin. 1958. Pp. 415. 30*s*.

In the last few years a spate of books on sexual perversion and in particular on homosexuality has risen on both sides of the Atlantic. In Britain this has served to supply the demand created by the publication of the Report of the Departmental Committee on Homosexuality and Prostitution, a body which owed its existence to the public and Parliamentary concern aroused by some sensational criminal prosecutions and convictions for homosexual offences. In this sense, Dr Allen's book can be regarded as, to some extent, a tract for the psychiatric times designed for the lay reader. Having briefly surveyed genetic and endo-

crinological theories of causation, the author expresses his view that homosexuality is essentially a psychological disorder, capable if taken in time of frequent amelioration or cure by various forms of psychotherapy with or without social treatment. He illustrates his contention by a number of abbreviated case histories describing apparent cures by psychiatric or social methods; the book ends with a summary and critique of the Wolfenden Report.

Written by a psychiatrist with lengthy experience in the diagnosis and treatment of homosexuality, Dr Allen's book will serve to offset somewhat the therapeutic nihilism regarding this condition which is all too common in psychiatric circles. It may also encourage those conflicted homosexuals who happen to read it to seek professional help or guidance—no mean service.

The volume edited by Dr Berg in Britain and in the U.S.A. by Mr Krich is of a different order. It is, in fact, an anthology of contributions to the subject by standard authorities, prefaced by a considered survey from the pen of Dr Berg and a section in which excerpts from the autobiographies, diaries and correspondence of homosexuals present some of the subjective aspects of the problem. This preliminary section incidentally, although appealing in content, can scarcely be regarded as representative of the reactions of that large group of 'non-conflicted' homosexuals who apparently live on terms with their perversion without any major disturbance of their mental function, or, at any rate, no more disturbance than could be found in matched heterosexual groups.

Owing largely to the discriminating selection of papers, the clinical section is unusually rewarding. Most omnibus books in which the articles are commissioned by the editor are extremely patchy in presentation and often end by confusing the reader. Dr Berg's selection, although bringing out differences in the theories and practices of the selected authors, provides a more homogeneous survey. Here are articles by, amongst others, Freud, Jung, Stekel, Ferenczi, Hirschfeld, Henry, Schilder and Bergler covering both the psychobiological and the psychoanalytical approach to the subject. This is an unusually successful venture which deserves a place on the bookshelves not only of the psychiatrist, but of all those who are interested in the individual and social aspects of an extremely complicated problem, one indeed that is largely unsolved and may indeed be insoluble save in its criminological aspects. For it is clear that the forces of social intolerance backed by outmoded legal codes contribute not a little to its refractoriness.

EDWARD GLOVER

Chronic Schizophrenia. By T. Freeman, J. L. Cameron and A. McGhie. London: Tavistock. 1958. Pp. x+158. 21*s*.

This monograph has three main components. The first is a clear, succint summarary, in 10,000 words or so, of the central concepts in the psychopathology of schizophrenia. Although reference is also made to Penfield, Tinbergen, and Lorenz, it draws heavily from Freud and later psychoanalysts, notably Hartmann and Katan, and is more narrowly psychoanalytic in view-point than are the second and third components, to which it serves as an introduction. The second is a description of the chief manifestations of schizophrenia, especially the reduced ability to distinguish self from not-self, the confusion of personal identity and the disorders in perception, thinking and memory. The third is a discussion of methods of treating chronic hospitalized schizophrenic patients, which are based upon an understanding of their difficulties and designed to restore ego-functions.

Following Federn, the authors argue that the patient's reduced capacity to experience himself as a separate and distinct entity, the loss of ego feeling, is the basic disturbance, and that the failure to distinguish internal and external sensations and the disorders in perception, thinking and memory are necessary consequences. The essential functions of the ego, with the associated 'secondary process', break down, they assert, and the 'primary process' is then released from the inhibitory influences of the ego. The emergence of the primary process accounts for many of the symptoms of schizophrenia, amongst other things, for the disturbance in associations and the 'splitting' in the thought processes, which Bleuler regarded as the basic symptoms. From this viewpoint, the authors describe observations which they have made in the course of a prolonged and intensive study of a sample of chronic patients. They do not here report their observations systematically as evidence which can be taken to prove their thesis. In the main, they report striking incidents anecdotally, which they then interpret in order to illustrate the thesis. So doing they make sense of much of the strange behaviour of patients which would otherwise appear unintelligible and go some way towards justifying their claim that, with all its limitations, the psychological view is in a position to furnish an integrated and consistent theory of chronic schizophrenic manifestations, and a theory which offers the possibilities of a rational, non-empirical therapy. All this they do very well. Yet they deliberately refrain from discussing aetiology and do no more than

hint at their views on the causes of the breakdown of ego functions. To regard this breakdown as basic, and other symptoms as consequences, allows a clear description of the disease, but whether this ordering of the symptoms is sound and useful can ultimately only be decided when the aetiology is specified.

In their discussion of treatment methods, the authors have kept closer to the psychopathology of the disease than have many of those who have recently written on the therapeutic community, although they cover several of the same points, such as greater freedom, the unlocking of doors, home visits and more meaningful occupation. Those readers who are unfamiliar with the conditions prevailing in the back wards of our mental hospitals may be astonished, indeed shocked, by some of the things which the authors need to say, but conditions to start with being no less primitive in their hospital than elsewhere, they had perhaps better opportunities to discover what are the essentials, and what are the frills. In their view the essentials lie in the patient's relationship with the nurse, and, in the early stages, in the fostering of primarary identifications particularly. They attach great importance to non-verbal methods. Much needs to be done, therefore, to revise the nurse's traditional role and to train and sustain her in a new role which requires from her a much greater understanding of her patients' needs.

This short monograph raises many important issues, which deserve further consideration along the lines suggested by its authors and perhaps also along more experimental lines. It is to be hoped that they will find opportunities to elaborate their ideas and then to report them more fully. In the meantime this monograph will serve well as a useful and refreshing introduction.

D. RUSSELL DAVIS

The Patient and the Mental Hospital. Ed. by M. GREENBLATT, D. J. LEVINSON and R. H. WILLIAMS. Glencoe, Illinois: The Free Press. 1957. Pp. xviii + 658. $ 6.00.

This large book, edited by a psychiatrist, a social psychologist and a sociologist, consists of a collection of papers by many different authors arranged under the headings: 'Mental Hospital Organisation', 'Therapeutic Personnel', 'The Ward' and 'The Patient and the Extra-Hospital World'. It continues, in somewhat the same way, the sort of approach of the now famous book *The Mental Hospital*, by Stanton and Schwarz; and, as might be expected, these two authors are among the contributors. Some of the chapters are bound to be a little embarrassing to those who work in mental hospitals. For example, Prof. Loeb's chapter on 'Role Definition in the Social World of a Psychiatric Hospital' endeavours to place psychiatrists, psychologists, social workers, nurses, occupational therapists, and so on, in some sort of hierarchial prestige order, and having done so finds that 'the peer of the chief nurse is someone low on the medical totem pole—low in the professional categories and departments above nurses'. It would be interesting to hear Prof. Loeb develop this theme to a group of British hospital matrons.

The organization of the larger State mental hospitals in America is so strikingly different from the way that the majority of British mental hospitals are run, that many of the comments made in this book have little relevance this side of the Atlantic. Nevertheless, much of interest is discussed, such as the rival claims of the hospital's maintenance and administration against its therapeutic activities; problems of status and authority, individuality and occupational label, communication and so on. Another interesting topic, on which some experimental work is quoted, is that of 'Custodialism' versus 'Humanism' in hospitals. A scale assessing these two attitudes correlates highly with other authoritarian ideological scales when applied to mental hospital staff.

Various experiments are reported where a team of sociologists have tried to change the staff attitude in some chronic wards, and to introduce new methods and a new atmosphere. As might be expected, striking changes in backward patients are not difficult to produce if the right steps are taken. The problem is to demonstrate that the new way is easier or more economic. Shortage of staff is as often the bar to such improvements as is the custodial attitude in the administration of the hospital. This is still almost a virgin field of study for social psychologists.

This book is by no means confined to self-conscious studies of interpersonal relations within the mental hospital community, and is as much concerned with the relationship of the patient to the outside world. It raises problems which we should all be thinking about, and discusses them in a manner refreshingly free from dogma.

RALPH HETHERINGTON

OTHER PUBLICATIONS RECEIVED

The Sociology of Knowledge: an Essay in Aid of a Deeper Understanding of the History of Ideas. By WERNER STARK. London: Routledge and Kegan Paul. 1958. Pp. 356. 36*s.*

The Door of Serenity: a Study in the Therapeutic Use of Symbol Painting. By AINSLIE MEARES. London: Faber and Faber. 1958. Pp. 119. 21*s.*

Educational and Other Aspects of the 1947 Scottish Mental Survey. Scottish Council for Research in Education. University of London Press. 1958. Pp. 148. 15*s.*

The Student-Physician: Introductory Studies in the Sociology of Medical Education. Ed. by R. K. MERTON, G. READER and P. L. KENDALL. Harvard University Press. London: Oxford University Press. 1958. Pp. 360. 40*s.*

Man and Time: Papers from the Eranos Yearbooks. Vol. 3. Ed. by J. CAMPBELL. London: Routledge and Kegan Paul. 1958. Pp. 414. 35*s.*

Freudianism and the Literary Mind (2nd ed.). By F. J. HOFFMAN. Louisiana State University Press. London: Oxford University Press. 1958. Pp. 350. 40*s.*

The Patient Speaks: Mother Story Verbatim in Psychoanalysis of Allergic Illness. By H. A. ABRAMSON. New York: Vantage Press. 1956. Pp. 239. $3.50.

The Human Hand. By GÉZA RÉVÉSZ. Tr. by JOHN COHEN. London: Routledge and Kegan Paul. 1958. Pp. 138. 21*s.*

Psychology and Religion: West and East. Volume eleven of the collected works of C. G. JUNG. London: Routledge and Kegan Paul. 1958. Pp. 699. 42*s.*

Widows and their Families. By PETER MARRIS. London: Routledge and Kegan Paul. 1958. Pp. 172. 18*s.*

New Ways in Management Training. By C. SOFER and G. HUTTON. London: Tavistock. 1958. Pp. 127. 15*s.*

The Undiscovered Self. By C. G. JUNG. Tr. by R. F. C. HULL. London: Routledge and Kegan Paul. 1958. Pp. 113. 10*s.* 6*d.*

The Circle of Guilt. By FREDRIC WERTHAM. London: Dennis Dobson. 1958. Pp. 211. 18*s.*

Your Speech Reveals Your Personality. By D. A. BARBARA. Illinois: Thomas. Oxford: Blackwell. 1958. Pp. 174. 42*s.*

Collected Papers: Through Pediatrics to Psychoanalysis. By D. W. WINNICOTT. London: Tavistock. 1958. Pp. 350. 35*s.*

The Matrix of Medicine: Some Social Aspects of Medical Practice. Ed. by N. MALLESON. London: Pitman Medical Publishing Co. 1958. Pp. 234. 45*s.*

Analyzing Psychotherapy. By S. KATZENELBOGEN. New York: Philosophical Library. 1958. Pp. 126. $3.00.

Home Conditions: a Sociomedical Study of 1066 Hospitalized Patients with Skin and Venereal Diseases. By E. LOMHOLT. Copenhagen: Rosenkilde and Bagger. 1958. Pp. 100. Dan.kr. 14.00.

Copywriting and its Presentation: a Guide to Advertisement and Sales Letter Writing. By F. W. JEFKINS. London: Crosby Lockwood. 1958. Pp. 238. 21*s.*

BINOCULAR RESOLUTION AND PERCEPTION OF RACE IN SOUTH AFRICA*

BY THOMAS F. PETTIGREW AND GORDON W. ALLPORT
Harvard University

AND ERIC O. BARNETT
University of Natal

Our experiment uses a newly developed method for research in social perception: a brief stereoscopic presentation of pairs of racial photographs. Members of the four major racial groups in South Africa—Europeans, Coloureds, Indians, and Africans—serve both as judges and as photographic stimulus material. The factor of eye dominance is controlled.

While the stimulus material itself accounts for many of the findings, there is clear evidence that 'directed state' factors are also at work. White South Africans, especially the Afrikaners, seem to manifest a 'perceptual vigilance' in their judgements. They report racially mixed photographs to be either pure *European* or full-blooded *African* more often than do judges of other races. This tendency is consistent with the bifurcation (judgements of 'black' or 'white') known to be characteristic of people who are apprehensive concerning race relations.

The use of meaningful stimuli presented in a setting of binocular rivalry offers new promise for the study of cultural and personal determinants in perception.

I. INTRODUCTION

The role of meaningful content in binocular rivalry and fusion has not yet received careful and systematic attention. E. Engel (1956) and J. Bagby (1957) have each recently shown, however, the lively possibilities that lie in this unexplored avenue of research. Engel has demonstrated, for instance, that *right-side up* faces predominate over simultaneous images of faces *upside down*. Employing more complex stimuli, Bagby used both Mexican and American subjects. Such pictures as a bullfight and a baseball game presented simultaneously sometimes give meaningful fusions of the two, but more often result in a conflict. The culturally less familiar picture tends to be suppressed, the culturally more familiar image tends to be seen.

Precisely what variables are operating to determine the part meaning plays in binocular resolution cannot be specified as yet. Familiarity is clearly one important factor in the studies of Engel and Bagby. It seems reasonable to suppose that stereoscopic viewing of meaningful material may offer likewise a sensitive technique for studying a wide variety of 'directed state' problems in perception. (We are using F. H. Allport's (1955) term 'directed state' to cover all psychogenic and psychodynamic influences on perception.) H. Cantril (1957) has suggested, for example, applications of the technique to psychiatric research.

* This study was conducted at the University of Natal, Durban, Union of South Africa, and received support from a grant to Princeton University for perceptual research from the Rockefeller Foundation.

For assistance as interpreters or in arranging and conducting the experimental sessions, we are indebted to Mr R. C. Albino, Mrs Louise Barnett, Dr W. Bodenstein, Mr H. J. Bosman, Mr H. Dickie-Clark, Mr J. Mann, Mr E. Mayisela, Mr C. Ramphal, Mr H. O. Wagner, Mr A. Williams, and Mr E. Zuma.

For valuable editorial and theoretical suggestions the writers wish also to express their appreciation to Dr H. Cantril, Dr E. Engel and Dr F. Kilpatrick.

The present paper uses the technique in a study of race relations. The general problem under investigation is to ascertain the principal variables related to identification of race. (We use the term 'race' throughout this paper to distinguish the four major ethnic groups of South Africa without claim to any biological or anthropological precision.) More specifically, we are interested in the 'directed state' question: *Do members of various ethnic groups identify each other differently?* This question is of special interest and importance in a highly diverse, multiracial society such as the Union of South Africa where the present study is set.

II. Method

(i) *Apparatus*

The research employed an Engel stereoscope which allows precise light and distance control. This apparatus is described in detail by Engel (1956).

The stimulus material consisted of similarly posed, full-faced, bust photographs of both sexes and of all four major ethnic groups of South Africa. These four groups are African (full-blooded Bantu), Coloured (mulattoes, frequently also possessing some Malayan ancestry, who form a legally and popularly recognized separate ethnic group), Indian (Hindu or Moslem), and European (white—either English or Afrikaans-speaking). This material was readily available from a local photographer who had on file a large number of passport photographs of all groups (about $1\frac{3}{4}$ in. square). These photographs are required by the South African government in connexion with their Population Registration Act. From this rich assortment we selected pictures of subjects between the ages of twenty-one and thirty years, these pictures being regarded as typical of each of the four groups by two white South African social scientists.

For our practice series, four duplicated bust photographs of one male from each of the four racial groups were employed. That is, identical pictures were shown to the two eyes. The experiment proper used forty bust pictures, twenty of each sex and ten of each race. Twenty same-sex pairs were formed so that each group would be paired with all four groups an equal number of times. In addition, these pairs were equated for 'pleasantness' of expression through the independent ratings of three judges. This equalizing of all pairs as to pose, age, sex, and pleasantness maximizes the likelihood of the fusion of binocular stimuli.

(ii) *Procedure*

Unlike Engel (1956), who placed no time limit on the exposure duration, we limited each exposure to approximately 2 sec. in order to minimize rivalry and encourage fusion. Our chief interest lay in impromptu racial identification. Pilot work indicated that the 2-sec. exposure was an optimal time for this purpose as longer intervals tended to induce uncertainty of judgement and fluctuation of images.

Though the Engel stereoscope allows for differential illumination of the two images, we maintained constant illumination for both eyes, using the maximum allowed by the 7 watt lamps.

The subject sat comfortably at a convenient height before the stereoscope. He was told that he would see pictures of people, some of them would be European, some Coloured, some Indian, and some African. At the signal 'Ready', the subject took the viewing position and with both eyes looked at the brief exposure of the two images. Immediately thereafter he named the race he thought he had seen represented, and if he made qualifications or seemed uncertain of his judgement the fact was noted. Actually, under these conditions we found that nearly all subjects made immediate and unequivocal judgements. The importance of this fact should be underlined. Whether the subject suppresses one image or creates a fusion of the two cannot always be known, but he promptly gives *some* resolution to retinal conflict on the basis of *subjective* factors.

For the preliminary practice series there was no time limit, and since the two photographs in each pair were identical, fusion was excellent and errors of naming rare. Given in a constant order, these practice trials served four functions. First, the subject 'warmed up' and became generally accustomed to the experimental procedure. Secondly, the subject was able to vary the distance of the pictures from his eyes; once he achieved maximum fusion at the optimal distance for him, this distance was maintained throughout the remainder of his experimental session. Thirdly, since he saw only the image of a single person (even if it were blurred or doubled in the process of adjusting the distance) this fact reinforced his belief that only one face was being presented to him each time. The fourth purpose of the practice

trials was to demonstrate that all four groups of South Africa—European, Coloured, Indian, and African—were represented in the pictures.

Immediately after the practice series, the experiment proper began. The following list gives the randomized order used (E = European, C = Coloured, I = Indian, A = African, m = male, and f = female):

(1) *If–Af*	(6) *Em–Cm*	(11) *Am–Am*	(16) *Em–Em*
(2) *Ef–If*	(7) *Ef–Af*	(12) *Cm–Am*	(17) *Em–Im*
(3) *Af–Af*	(8) *Ef–Ef*	(13) *Cf–Cf*	(18) *Em–Am*
(4) *Ef–Cf*	(9) *Cf–Af*	(14) *Cm–Im*	(19) *If–If*
(5) *Cf–If*	(10) *Im–Am*	(15) *Cm–Cm*	(20) *Im–Im*

The picture listed first in each pair was presented to the right eye while the other photograph was presented to the left eye. Upon finishing the exposure of these twenty pairs, the list was reversed in order to control for eye dominance. Thus, with the twentieth pair (*Im–Im*) the same Indian male would appear in the twentieth trial before the left eye and in the twenty-first trial before the right eye. The twenty-second trial employed the same pair as the nineteenth trial but with the pictures reversed as to eyes, and so on until the fortieth exposure.

At the close of each experimental session, the investigator determined each subject's educational background, occupation, and age.

(iii) *The subjects*

One-hundred and twenty-two subjects from the Durban area representing five different ethnic groups participated in the study. The data from fourteen additional subjects were discarded by a previously chosen criterion of extreme eye dominance. In all of these rejected cases, at least eight out of the twelve two-race pairs (nos. 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 17 and 18) were consistently responded to by the label of the race represented to one particular eye. Thus, a right-eye dominant person would report with the first pair an Indian when the Indian was shown the right eye on the first trial and then report an African when the African was shown the right eye on the fortieth trial. When this happened eight or more times with the same eye the subject was not included in the final sample. The remaining 122 subjects all possessed normal eyesight in other respects as well as showing no severe eye dominance.

Many of the African and Coloured sample members of lower economic status received a shilling for their services. Other subjects were not paid.

The distribution of subjects is as follows. European Afrikaans-speaking, $N = 20$: European English-speaking, $N = 22$: Coloureds, $N = 20$: Indians, $N = 20$: Africans, $N = 40$. For an understanding of our results it is important to note that the European subjects comprise two sub-groups: Afrikaans-speaking and English-speaking.

The subjects were residents of Durban, excepting for the Coloureds who lived in a village south of the city. The average age for all groups was in the middle twenties. Slightly less than half of each ethnic group were women. The educational range for the most part covered medium and low amounts of schooling; very few had matriculated in college. The possible influence of the educational variable will be considered later. The Africans were servants excepting for a few student nurses, while the Indians and Coloureds for the most part held lower skilled jobs; the European subjects were in middle class occupations. In spite of our efforts to obtain a homogeneous socio-economic group of subjects, the South African social setting makes such equating impossible to achieve.

Obviously no claim can be made that these samples were randomly selected and therefore representative of the various ethnic groups of the Union of South Africa. L. Kish (1957) has noted that such clustered samples frequently yield misleading probability statements when simple random sample statistics are used. Hence probability statements made in this article should be viewed as tentative.

Social conditions in South Africa today make it impossible for a single experimenter of any race to collect data from the various groups. Accordingly, the procedure was standardized as much as possible and various investigators tested different groups. We are indebted to Messrs Mayisela and Zuma for collecting data from African subjects, to Mr Ramphal for Indian subjects, and to Messrs Mann and Dickie-Clark for Coloured subjects.

III. THE GUIDING HYPOTHESIS

The exploratory nature of the present study precluded the use of specific hypotheses. One general hypothesis did guide the work, however: *members of different ethnic groups will vary in their racial identifications of each other.* Put differently, identification of race

is in part a function of a person's own race. In order to examine this hypothesized relationship further, the study varied the type of stimulus, using one-race pairs of pictures and all possible mixtures of two race pairs of photographs.

The special interest of the study focuses upon the European groups—particularly the Afrikaans-speaking Europeans. The reason, of course, is that at the present time the white people in South Africa are living in a state of considerable apprehension and are endeavouring to maintain white supremacy at all costs. According to the 'directed state' theories, such emotional tension might lead to either 'perceptual vigilance' (in which case an excessive number of the threatening races might be perceived) or to 'perceptual defence' (in which case an unusually small number of the threatening races might be perceived). A related hypothesis would hold that Europeans being under stress would tend to show excessive bifurcation of judgements. That is, they would judge too many photographs to be either European or African (white or black) and too few as Indian or Coloured.

Our attention was drawn to these possibilities through a small pilot study involving well-educated Afrikaans-speaking and English-speaking South African whites. While both groups undoubtedly are under strain owing to the racial problems of South Africa, the Afrikaners gave quite different judgements from the English-speaking. For example, while looking at a two-race stimulus involving both European and Indian pictures, the Afrikaners uniformly reported *European* (apparently suppressing the Indian image) while the English reported *Indian* or *Coloured* frequently. This result we tentatively ascribed to the Afrikaner's greater racial apprehensiveness.

In short, 'directed state' possibilities with the various groups of subjects—and especially the Afrikaners—are given particular attention and emphasis in the present study.

IV. Results

(i) *General overview*

Table 1 presents the entire raw data obtained in this investigation, from which all later analyses will be derived. For purposes of illustration, let us read the first line. It shows that all eighty judgements of the European-European stimulus combination made by the twenty Afrikaner subjects (four judgements per subject) were *European*. For the twenty-two English-speaking subjects, the European-European combination was judged

Table 1. *Racial identifications of five observer groups in response to ten stimulus combinations**

Binocular stimulus combinations (four of each)...	Afrikaners $N = 20$				English $N = 22$				Coloureds $N = 20$				Indians $N = 20$				Africans $N = 40$		
	E	*C*	*I*	*A*	*E*	*C*	*I*	*A*	*E*	*C*	*I*	*A*	*E*	*C*	*I*	*A*	*E*	*C*	*I*
1. European–European	80	—	—	—	82	—	6	—	52	15	13	—	71	1	8	—	108	14	36
2. Coloured–Coloured	3	47	20	10	—	69	18	1	2	64	14	—	4	57	19	—	5	104	47
3. Indian–Indian	—	9	65	6	2	14	72	—	4	21	55	—	—	10	70	—	5	36	118
4. African–African	—	3	—	77	—	3	2	83	—	—	1	79	—	2	—	78	1	7	5
5. European–Coloured	34	28	11	7	32	45	5	6	32	35	12	1	21	52	5	2	41	73	35
6. European–Indian	43	7	29	1	38	10	40	—	37	3	39	1	20	1	59	—	52	11	97
7. European–African	20	18	14	28	28	20	18	22	16	21	23	20	18	28	19	15	31	38	49
8. Coloured–Indian	4	19	52	5	—	37	47	4	1	18	61	—	2	26	52	—	12	47	100
9. Coloured–African	—	22	7	51	—	34	10	44	—	34	6	40	—	37	2	41	—	76	7
10. Indian–African	—	4	47	29	—	5	60	23	—	5	57	18	—	9	41	30	—	27	81
	184	157	245	214	182	237	278	183	144	216	281	159	136	223	275	166	255	433	575

* *E*, *C*, *I*, and *A* refer to the responses made by the subject: *E* = European, *C* = Coloured, *I* = Indian, and *A* = African.

eighty-two times as *European* and six times as *Indian*. Continuing across the first line, we note the Coloured subjects saw the European-European pair as being *European* fifty-two times, *Coloured* fifteen times, and *Indian* thirteen times; while the Indian subjects saw the pair as *European* seventy-one times, *Coloured* once, and *Indian* eight times. Completing the line, the stimulus pair was seen by the forty African subjects as *European* 108 times, as *Coloured* fourteen times, as *Indian* thirty-six times, and as *African* twice.

(ii) *Racial identifications*

Since in fact the series of photographs employed an equal number of European, Coloured, Indian, and African, we can first ask, Do the judgements fall equally—25% for each group? Table 2 presents the percentages bearing on this question. For the one-race pairs (stimulus combinations 1–4) there was a strong tendency for the judgements to vary from an equal usage, 25% for each group ($2 \times 4\ \chi^2 = 23{\cdot}50$; $p < 0{\cdot}001$). For these stimuli, *European* identifications were disproportionately underrepresented and *Indian* identifications disproportionately overrepresented.

Table 2. *Percentage of responses assigned to each group by all subjects combined*

	Judgements			
	European	Coloured	Indian	African
One-race pairs (combinations 1–4)	21·5	24·4	29·1	25·0
Two-race pairs (combinations 5–10)	16·5	27·0	37·1	19·5
Total (combinations 1–10)	18·5	25·9	33·9	21·7

For the two-race pairs (stimulus combinations 5–10) the tendency to vary from an equal distribution of 25% was even more pronounced ($2 \times 4\ \chi^2 = 300{\cdot}62$; $p < 0{\cdot}001$). Judgements of *Coloured* and *Indian* were more numerous than might be expected, judgements of *European* and *African* less so. This same trend holds for all ten stimulus combinations taken together ($2 \times 4\ \chi^2 = 260{\cdot}78$; $p < 0{\cdot}001$).

Summing up, the four racial identifications were not employed with equal frequency. The tendency for *Coloured* and *Indian* judgements to be more common than *European* and *African* judgements was pronounced for the more ambiguous stimuli—the two-race pairs.

Table 3. *Percentage of judgements assigned to each group by each subject group*

	Judgements			
Subjects	European	Coloured	Indian	African
Afrikaners	23·0	19·6	30·6	26·8
English	20·7	26·9	31·6	20·8
Coloured	18·0	27·0	35·1	19·9
Indian	17·0	27·9	34·4	20·7
African	15·9	27·1	35·9	21·1

The differences noted in Table 2 in the differential use of the four race labels raises the further question concerning the five groups of subjects. Do all groups of subjects evidence the trend noted in Table 2 or are there differences between the groups in the frequency with which they use the various labels? The relevant data for this question are summarized in Table 3.

All groups save the Afrikaners reveal the same general trend shown in Table 2: the

more frequent use of *Coloured* and *Indian* labels than *European* and *African* labels. Our group of greatest interest, the Afrikaners, however, reports more *Europeans* and *Africans* and less *Coloureds* and *Indians* than any other subject group. Though much less than the Afrikaners ($4 \times 2\ \chi^2 = 16{\cdot}80$; $p < 0{\cdot}001$), the English participants evidence some tendencies in the same direction. Only chance differences exist between the Coloured, Indian, and African subjects, all of whom show a strong tendency to over-use *Coloured* and *Indian* labels.

A glance back at Table 1 will show the reader that these differences between Afrikaners and other subjects in the use of racial labels are quite general in nature and not limited to a few isolated stimulus combinations. With all ten combinations, Afrikaans-speaking subjects tended with minor exceptions to see more *Europeans* and *Africans* than did the other four subject groups.

(iii) *The accuracy of judgements*

How accurate were the judgements of race for the pairs of pictures involving only one race? The reader will recall that there were eight pairs of such pictures, one for each sex for all four races (combinations 1 through 4). Each of these pairs having been shown twice, there were sixteen judgements made by each subject to stimuli which realistically represented only one race. The accuracy of these judgements is summarized in Table 4.

Table 4. *Percentage of correct identifications of one-race pairs*

Subject groups		The one-race pairs				Total average
	N	European	Coloured	Indian	African	
Afrikaners	20	100·00	58·75	81·25	96·25	84·06
English	22	93·18	78·40	81·83	94·33	86·94
Coloureds	20	65·00	80·00	68·75	98·75	78·13
Indians	20	88·75	71·25	87·50	97·50	86·25
Africans	40	67·50	65·00	73·75	91·88	74·53
Total average		80·53	69·88	77·87	95·08	80·84

Three interesting questions can be asked concerning accuracy. First, which races were most accurately perceived? Table 4 shows that the African pairs were by far the easiest to identify correctly (African-European corrected $2 \times 2\ \chi^2 = 46{\cdot}92$; $p < 0{\cdot}001$), while the Europeans and Indians were of average difficulty (European-Indian corrected $2 \times 2\ \chi^2 = 0{\cdot}89$; not significant). The Coloureds were the hardest to identify correctly (Indian-Coloured corrected $2 \times 2\ \chi^2 = 10{\cdot}14$; $p < 0{\cdot}01$).

Secondly, were some groups of subjects more accurate in their judgements than others? Immediately we notice that, except for the Africans, each race recognizes itself better than do others. Thus, Afrikaners and English are the most correct judges of the European pairs, Coloureds of the Coloured pairs, and Indians of the Indian pairs. Even relatively inaccurate Africans were at their best in judging the African pairs. We note in Table 4 sharp differences between the relatively inaccurate Coloureds and Africans on the one hand and the relatively accurate Afrikaners, English, and Indians on the other. The reader will remember, though, that the Africans and Coloureds were comparatively the poorest educated groups of our five samples. If education is an important factor in correct identifications, then Africans and Coloureds would have been expected to perform less accurately for educational reasons alone. Thus, we shall need to check later on the effects of education in this relationship.

A third question concerns the 19% incorrect identifications of the one-race pairs. How were these errors distributed among the four labels? Table 5 presents the pertinent data.

We note from Table 5 that when errors occurred in response to the one-race pairs they predominantly consisted of mistakenly labelling some race as *Coloured* or *Indian*. The two most accurately perceived races, African and European (see Table 4), were seldom guessed incorrectly. Only one exception to this trend occurs—the Afrikaans-speaking subjects again. Of their fifty-one errors to the same race pairs, the Afrikaners reported *African* far more often than all other groups combined and guessed *Coloured* and *Indian* less than any other subject group. A further check on our master chart, Table 1, reveals that these differences are largely a result of the Afrikaners' differential responding to stimulus combinations 2 and 3—the Indian and Coloured one-race pairs. These unpopular groups are seen as 'black'. Sixteen of the twenty-two times these pairs were mistakenly seen as African occurred with Afrikaans-speaking perceivers. This is a special case of the tendency noted earlier for the Afrikaners to see less Coloureds and Indians and more Africans than other groups.

Table 5. *Distribution of errors for mistaken one-race pairs*

Subject groups	Total errors	Percentage of inaccurate responses			
		European	Coloured	Indian	African
Afrikaners	51	5·9	23·5	39·2	31·4
English	46	4·3	37·0	56·5	2·2
Coloureds	70	8·6	51·4	40·0	0·0
Indians	44	9·1	29·5	61·4	0·0
Africans	163	6·7	35·0	54·0	4·3
Totals	374	7·0	36·1	50·5	6·4

Table 5 also shows interesting group differences in incorrectly judging non-European pictures as *Europeans*. The two white groups committed this error less than any of the three non-white groups. Also we note that the Coloureds and Indians tended to perceive their own race more than other subjects in their incorrect judgements.

(iv) *The perception of the two-race pairs*

In the previous section we have considered the judgements made of the one-race stimulus combinations. Let us now review the results of the six two-race stimulus combinations (5–10), the more ambiguous stimuli which do not represent any one race.

Table 6 summarizes the responses made to the ambiguous pairs. A 'suppression' is defined as the naming of *either* of the two races represented in the two-race pair. With the European-African stimulus combination, for instance, either *European* or *African* is scored as a 'suppression' response. A 'suppression' response means therefore that one of the two alternatives is selected for report. By contrast, a 'manifest fusion' response occurs when the observer reports a third (usually intermediate) race. For example, with the European-African combination, he reports either *Coloured* or *Indian*.

Table 6 reveals very little difference between subject groups in their overall responding to these two-race pairs of photographs. If, however, we examine within the various stimulus combinations we note several points of interest. The most striking of these occurs with the European-African combination; slightly over half of the responses made to this stimulus combination were manifest fusions—in other words, were judgements of *Coloured*

or *Indian*. This figure contrasts with the less than 20% manifest fusions for each of the other five two-race pairs. We note, however, that the European subjects—particularly the Afrikaners—gave more suppression judgements than did the non-European subjects in the case of the European-African mixture. Thus, as Table 6 shows, the Afrikaners gave 60% bifurcated (black or white) judgements as compared, say, with the Indian subjects who gave only 37·5% of such judgements.

The Afrikaners also vary from the other groups in their responses to the Coloured-Indian two-race pair. They tended to make fewer suppression judgements and consequently more *European* and *African* judgements to this stimulus combination than other subjects.

Table 6. *Percentage of 'suppression' identifications of ambiguous, two-race pairs**

Subject groups	*N*	Stimulus combinations: *E–C* (5)	*E–I* (6)	*E–A* (7)	*C–I* (8)	*C–A* (9)	*I–A* (10)	Total average
Afrikaners	20	77·5	90·0	60·0	88·8	91·3	95·0	83·8
English	22	87·5	88·6	56·8	95·5	88·6	94·3	85·2
Coloureds	20	83·8	86·4	45·0	98·8	92·5	93·8	83·4
Indians	20	91·3	98·8	37·5	97·5	97·5	88·8	85·2
Africans	40	71·3	93·1	45·6	91·9	95·6	83·1	80·1
Total average		80·5	93·0	49·2	94·1	93·4	89·8	83·3

* 'Suppression' is defined as a response of *either* of the two races represented in the mixed-race pair. For example, with the European–Coloured stimulus combination, either European or Coloured would be scored as a suppression response.

Table 7. *Distribution of manifest fusions of two-race pairs**

Subject group	Total manifest fusions	Percentage manifest fusions: European	Coloured	Indian	African
Afrikaners	78	5·1	37·2	41·0	16·7
English	78	0·0	44·9	42·3	12·8
Coloureds	73	1·4	39·7	56·2	2·7
Indians	68	2·9	55·9	38·2	2·9
Africans	191	6·3	39·8	47·6	6·3
Totals	488	3·9	42·4	45·7	8·0

* 'Manifest fusion' is defined as a response which names *neither* of the two races represented in a particular mixed-race pair of pictures. For example, with the European–Coloured stimulus combination, either Indian or African would be scored as a manifest fusion.

Additional evidence for this bifurcated tendency is seen in Table 7. One would normally expect manifest fusions (the naming of some race not represented in the stimulus pictures) to yield a judgement of *Coloured*—since this is the only legally designated mixed race in South Africa; or perhaps the judgement of *Indian* since in fact Indian photographs partake of dark complexions as well as European features of face and hair. All groups do in fact show this expected trend; 88% of the total manifest fusions are reported as either Coloured or Indian. At the same time, the Afrikaners, in whom we are especially interested, manifest this tendency less than other groups. Table 7 shows that they tend to report *Europeans* or *Africans* more often (22%) and *Coloureds* or *Indians* less often (78%) than other groups. We note also that the English-speaking subjects tend to resemble the Afrikaans-speaking subjects in their relatively large number of African manifest fusions (possibly evidence for perceptual vigilance).

(v) *Education and judgements of race*

We mentioned previously that the Coloured and African subjects were on the average somewhat less educated than the other three groups. When it was found that the Coloured and African subjects were also the least accurate in judging the one-race pairs (Table 4), the question was raised as to the relationship between education and accuracy. For the nineteen relatively well educated European subjects (eighth standard or higher) 86·2% of the one-race stimuli were judged correctly, while the six poorly educated European subjects (sixth standard or less) had 88·5% of their judgements correct. Though no educational effects were apparent in these European data, non-white subjects did evidence a very slight positive relationship. Seventy-nine per cent of the nineteen well educated non-European subjects' one-race judgements were accurate as compared to 77·3% of the judgements made by the poorly educated non-white subjects. Education seems only slightly if at all related to one-race accuracy in our stereoscopic situation.

There are still further indications in the data that the degree of education is not a crucial variable. Our special group of interest, the Afrikaners, revealed a number of interesting differences in their racial identifications from the English and Indian subjects, yet all three of these subject groups were fairly well educated. The unique results of the Afrikaans-speaking participants cannot be accounted for in terms of educational differences.

One difference did occur between the relatively well educated and the poorly educated subjects, both European and non-European. The better educated respondents tended to make more *Coloured* identifications than the relatively uneducated subjects. Twenty-six per cent and 29·4% of the higher educated European and non-European responses were *Coloured* as compared with the 21% and 25% figures of the poorly educated. The South African legal definition of a Coloured being any individual of mixed race means that those of our sample who were relatively well educated were in fact responding more accurately in a legal sense to the two-race stimulus combinations.

V. DISCUSSION

The use of meaningful stimuli in stereoscopic research presents many complexities. Photographs, particularly if not expertly done, are often slightly 'muddy' in appearance. We are inclined to believe that this source of error, for example, partly accounts for the unusually large number of *Indian* judgements made by our subjects. Moreover, limitations are introduced by tachistoscopic presentation. Two-second exposure such as we used makes it impossible to study the phenomena of rivalry and fluctuation as Engel (1956) succeeded in doing with longer intervals of presentation. An additional complexity is created by eye dominance. Unless adequate controls can be employed, the mere physical fact of eye dominance will seriously affect the results.

While our rapid exposure method limits the range of problems that can be studied, it does have certain advantages. We employed it in this study both to minimize fluctuation and to simplify communication with our subjects, with half of whom an interpreter was necessary.

Two methods controlled the eye dominance factor: (1) reversal of the order of

presentation of the stimuli to the right and left eyes, and (2) discarding the data of the 10% of our subjects who showed severe eye dominance.

In spite of the difficulties of stereoscopic research of this sort, we believe we have certain important findings. The present data clearly show that identification of race is a function of both the stimuli being perceived and the perceiver making the judgements. Let us consider each in order.

(i) *The stimulus as a factor in racial identification*

Recent studies in social perception have stressed perceiver variables almost to the exclusion of stimulus factors. Our data, however, indicate that the largest differences in racial identification are due to differences in the stimuli. First, it should be remembered that our entire subject population of 122 persons is correct in judging the one-race pairs on the average of four out of five times. For reasonably unambiguous stimuli, then, the racial identifications of our subjects are generally accurate (Table 4).

Another important point concerns the discriminability of the four races employed in this investigation. As represented by our photographic stimuli, the four groups are differentially difficult to judge correctly. Table 4 indicates that the one-race pairs of Africans are named nineteen out of every twenty times, and the one-race pairs of Europeans are correctly named eight out of every ten times. Indians and Coloureds, particularly the latter, are more difficult to judge accurately.

The two most accurately perceived groups—Africans and Europeans—are the most distinctly different in appearance of the races used. Table 6 shows that the stimulus combination of African and European is judged in terms of fusion far more than any other two-race combination. That is, over half of the judgements made of the African-European mixture are either *Coloured* or *Indian*; all other stimulus combinations involve naming at least four out of every five times, one of the two races actually represented. Such racial cues as skin colour, hair texture, and facial features are apparently in great conflict in the European-African combination and the conflict tends to be resolved with an intermediate judgement involving a 'mixed-cue' race. Europeans and Africans can be thought of as polar-opposite 'anchors' for the racial identifications of our subjects.

Indians and Coloureds are intermediate, mixed-cue races. They are the most difficult to judge, blending as they do features of both of the anchoring races. They are also the most frequently employed racial labels—particularly in response to ambiguous stimuli. Table 2 shows the disproportionately large use of the *Indian* and *Coloured* when judging two-race pairs; 64% of all judgements being either *Indian* or *Coloured*. Of the errors made in response to the one-race pairs, Table 5 shows that 87% are either *Indian* or *Coloured*. Moreover, Table 7 reveals that when fusion of the two images is manifest in the response to two-race stimuli, 88% of these fusions are either *Coloured* or *Indian*.

An additional point should be made concerning the greater frequency of *Indian* and *Coloured* judgements. By South African law, all racial mixtures are Coloured. The legally accurate response to all of the two-race stimuli should have been, therefore, *Coloured*. The fact that more *Indian* than *Coloured* judgements were made (Table 2) suggests that other cues—not legal definitions—are of importance in determining racial perception. Both the unusually frequent occurrence of *Indian* judgements (Table 2) and the more accurate judging of Indian as compared to Coloured one-race pairs (Table 4) may be due in part to the fact that Durban, the locale of this study, contains roughly eight Indians

to every one Coloured. Due to greater familiarity with Indians, our subjects might have tended more often to resolve a conflicting cue stimulus as an *Indian*, while persons in Cape Town—a city with many more Coloureds than Indians—might well have resolved the same conflicting cue situation by judging the stimulus mixture to be *Coloured*. Our present data do not allow us to estimate the importance of this factor of familiarity.

Summarizing, the racial identifications of our subjects varied considerably as a function of the stimuli. The European and African stimuli are more discriminable: these races are most dissimilar in appearance and social status; they seem to form polar-opposite anchorage points for the identifications. The Indian and Coloured stimuli are more difficult to identify accurately. Moreover, *Indian* and *Coloured* judgements are more often employed in response to two-race stimuli and more often used when errors occur in judging one-race pairs. The Indian and Coloured stimuli, then, seem intermediate between the European and African anchors. Taken together, all of this evidence indicates that our subjects' racial identifications are in large part determined by the particular stimuli being judged.

(ii) *The perceiver as a factor in racial identifications*

Factors of directed state likewise play a part. Our guiding hypothesis predicted that members of different ethnic groups would vary in their identifications of each other. The possible operation of perceptual defence, perceptual vigilance, and bifurcation of judgements was expected.

Throughout the results, the Afrikaners deviate most consistently in their judgements from the other four groups. They tend to respond in an either-or bifurcated manner. Let us review the evidence.

Table 2 shows the tendency for our subjects to give more *Coloured* and *Indian* than *European* and *African* judgements—especially when responding to the ambiguous, two-race pairs. Afrikaners, however, differ from the other subjects in this respect. In Table 3 we note that they give the smallest percentages of *Coloured* and *Indian* and the largest percentages of *European* and *African* judgements of any group. (In giving their judgements Afrikaners refer to Africans as *Bantu*, *Native*, or *kaffir*. Part of the reason for these less courteous labels is that in the Afrikaans language the term *African* would be *Afrikaner*!)

This bifurcated (black-white) responding of Afrikaners is further illustrated in their perception of one-race pairs. While moderately accurate in recognizing Indian and African one-race pairs, Afrikaners are the most accurate in judging European one-race pairs and the least accurate in judging the Coloured one-race pairs (Table 4). Even more interesting is the Afrikaners' distribution of errors for mistaken one-race pairs (Table 5). When they incorrectly identify one-race pairs, they make less *Coloured* and *Indian* judgements than any other group, and twice as many *African* judgements as all other groups combined. In short, Afrikaners tend to be comparatively accurate in judging one extreme—the Europeans, while frequently employing the other extreme—the Africans—as the judgement for all other one-race pairs.

The responses made to the two-race stimuli by the Afrikaans-speaking subjects furnish additional evidence for the bifurcation tendency. When judging the European-African mixtures, the Afrikaners tend to suppress one of the images and report either *European* or *African* more than any other group (Table 6). Similarly, when viewing the two-race pairs involving the two 'intermediate' races, Afrikaners again give more European and African judgements than the other four subject groups.

Table 7 provides further analysis of the viewing of the two-race pairs. In their manifest fusions (responding with a race not represented in the mixed pair), Afrikaners have a larger percentage of *African* and a smaller percentage of *Coloured* judgements than any other group.

The black-white, bifurcating judgements of our Afrikaner subjects were evident, then, in response to all of our stimulus combinations, as a glance back to Table 1 will confirm. Previous work in this area suggests several possible reasons for our results.

Some investigators have noted a bifurcating tendency in the judgements of highly involved subjects. C. Hovland and M. Sherif (1952), working with the sorting of items for a Thurstone-type Negro scale, found that 'Negro subjects and militantly pro-Negro white subjects tended to place a disproportionate number of statements in the extreme categories, in line with the predictions from judgment theory that individuals with strong attitudes tend to see issues in "black and white" and displace neutral statements to the extremes'. Later, Sherif & Hovland (1953) noted that subjects with instructions to use any number of categories they wanted tended to employ fewer categories in direct proportion to increased involvement in the statements being rated. In line with the bifurcation hypothesis, highly involved subjects generally require a minimum number of categories and ignore small differences between stimuli.

Additional interest has been generated in bifurcated judgements by the work of Adorno, Frenkel-Brunswik, Levinson & Sanford (1950) on the authoritarian personality. E. Frenkel-Brunswik (1949) first reported that authoritarians typically were 'intolerant of ambiguity' and responded in an either-or fashion. For highly authoritarian people, she writes, 'a simple, firm, often stereotypical, cognitive structure is required. There is no place for ambivalence or ambiguities. Every attempt is made to eliminate them, but they remain as potentials which might interfere at any time. In the course of these attempts a subtle but profound distortion of reality has to take place, precipitated by the fact that stereotypical categorizations can never do justice to all aspects of reality' (Adorno *et al.* 1950, p. 480).

Considerable evidence supports these views. Authoritarians disproportionately often agree with such categorical items as 'there is only one right way to do anything' and 'there are only two kinds of people: the weak and the strong' (Adorno *et al.* 1950). E. Roper (1946) found in analysing public opinion poll results that respondents relatively high in anti-Semitism typically gave either a 'yes' or a 'no' but rarely a 'don't know' answer when asked their opinion on current events. The racially apprehensive person apparently needs to give a definite response, not an ambiguous 'don't know'. Further support for this tendency has been summarized by G. W. Allport (1954, pp. 400–3).

It is difficult to ascertain from our data the predisposing factor underlying the bifurcating tendencies of our Afrikaner subjects. It seems reasonable, however, to assume that these results reflect the Afrikaners' heightened concern for, and deep involvement in, the racial issues of the Union of South Africa. Like Hovland and Sherif's subjects, they may be yielding black-white, either-or judgements more than other groups because of their greater involvement. The authoritarian personality research suggests the possibility that special personality factors may also have been operating to cause the bifurcating results noted in our data. More detailed study is required to assess these various possibilities.

Our data offer further indications that racial identifications are affected by the perceiver's own race. In Table 4, for instance, we note that for three out of the four races

one-race pairs are most accurately judged by the racial group represented. Afrikaans-speaking and English-speaking Europeans are the most correct in judging European pictures, Coloureds in judging Coloured pictures, and Indians in judging Indian pictures. Even the one exception, Africans, are more accurate in recognizing African pairs than they are in recognizing any other one-race pairs. Such a finding seems plausibly explained in terms of greater familiarity with one's own ethnic group.

Table 4 also shows that Coloureds and Africans are considerably less accurate in judging one-race pairs generally than are the other groups. Since education is not strongly related to accuracy in our data, exceptionally restricted contact with other races by our Coloured and African samples may be one possible reason for this finding.

Another group-related difference is seen in Table 5. When errors occur in judging one-race pairs, Coloured and Indian subjects tend to report their own race too often. By way of contrast with this hospitable orientation, the two European groups are most careful *not* to judge non-European pairs as European. This result must be partly a function of the difference in discriminability of the races noted in the previous section; Europeans being easier as a race to identify correctly than either Coloureds or Indians. But in addition, differential status between the races may also be a factor. As the high status groups, the Afrikaners and English have the most to lose by including others in their group and hence are the most discriminating concerning the use of the European label. Coloureds and Indians, on the other hand, being of lower status, are more willing to judge each other's pictures or European pictures as belonging to their own group. A glance back to Table 1 shows us that our Coloured and Indian subjects sometimes judge the European one-race pairs as their own race, but never judge their own one-race pairs as African—the lowest status group of all.

A final relationship concerns education and the number of *Coloured* responses. As has been mentioned, all of the two-race mixtures should legally be designated as Coloured. Thus, the tendency for the relatively well educated subjects—both European and non-European—to give more *Coloured* responses is in the direction of technical accuracy. The difference, therefore, may reflect the wider knowledge better educated people have concerning the legal situation.

It seems clear from a review of these findings that our general hypothesis is supported: racial identifications are determined in part by subjective variables and especially by the subject's own racial membership.

VI. CONCLUSIONS

The use of meaningful pictorial stimuli in a setting of binocular rivalry proves to be a promising technique for exploring the determinants of 'social perception', especially those sets often called *defence* and *vigilance*. In this paper we offer no particular theory of cognition to account for the results, but are content to establish the fact that cultural memberships, expectancies, and subjective attitudes (directed states) play some part in the resolution of binocular conflict. Important for any theory is the rapidity with which the resolution occurs employing both objective and subjective cues. In less than 2 sec. our subjects reach a decision with high subjective certainty concerning which one of four races is represented by the binocular stimuli, even in those cases where two *different* races appear before the two eyes. There is no awareness of binocular conflict; a single judgement is promptly given. While we cannot know the precise course of the complex

weighting of objective and subjective cues, nor the exact process by which fusion (or suppression) takes place, we nevertheless clearly demonstrate the presence of what M. D. Vernon (1957) calls 'cognitive inference in perceptual activity'.

When both eyes receive the image of two different members of the same race (and sex) the great majority of subjects give correct identifications. The accuracy is greatest for the two most discriminable groups—the Europeans and the full-blooded Africans. When different races are presented to the two eyes they are seen most often as 'intermediate', Coloureds or Indians. But it is in the departure from this modal tendency (veridicality) that the subjective determinants are evident. In general, observers are most accurate in judging members of their own group; and we show that familiarity with any group has favourable effects on identification. When errors are made, they are often of such an order as to preserve the status of one's own group.

Our most striking evidence for the role of directed states in binocular resolution comes from the tendency of Afrikaans-speaking whites to sort their judgements into two categories—white and black. While the number of our Afrikaner subjects is small, the trend finds some support from the similar but less pronounced results among English-speaking South African whites. Broadly speaking both groups feel threatened by the non-European majority, the largest segment of which is the African Bantu. To maintain white supremacy it seems functionally useful for Europeans to regard all non-Europeans as 'blacks', since in this direction the 'menace' lies. Intermediate perceptions would, so to speak, simply distract one from the main racial threat. In view of the more intense convictions of Afrikaners (on the average) in favour of white supremacy it comes as no surprise that their tendency toward perceptual bifurcation is more marked than among the English-speaking South Africans who, on the whole, are ideologically less tense.

References

Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J. & Sanford, R. N. (1950). *The Authoritarian Personality.* New York: Harper.

Allport, F. H. (1955). *Theories of Perception and the Concept of Structure.* New York: Wiley.

Allport, G. W. (1954). *The Nature of Prejudice.* Cambridge, Mass.: Addison-Wesley.

Bagby, J. (1957). A cross-cultural study of perceptual predominance in binocular rivalry. *J. Abnorm. Soc. Psychol.* **54**, 331–4.

Cantril, H. (1957). Perception and interpersonal relations. *Amer. J. Psychiatry*, **114**, 119–26.

Engel, E. (1956). The role of content in binocular resolution. *Amer. J. Psychol.* **69**, 87–91.

Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personal.* **18**, 108–43.

Hovland, C. I. & Sherif, M. (1952). Judgmental phenomena and scales of attitude measurement: item displacement in Thurstone scales. *J. Abnorm. Soc. Psychol.* **47**, 822–32.

Kish, L. (1957). Confidence intervals for clustered samples. *Amer. Sociol. Rev.* **22**, 154–65.

Roper, E. (1946). United States anti-Semites. *Fortune*, 257 ff.

Sherif, M. & Hovland, C. I. (1953). Judgmental phenomena and scales of attitude measurement: placement of items with individual choice of number of categories. *J. Abnorm. Soc. Psychol.* **48**, 135–41.

Vernon, M. D. (1957). Cognitive inference in perceptual activity. *Brit. J. Psychol.* **48**, 35–47.

(*Manuscript received* 26 *August* 1957)

ATTENTION TO VISUAL DISPLAYS DURING A VIGILANCE TASK

I. BIASING ATTENTION

By C. H. BAKER

Applied Psychology Research Unit, Medical Research Council, Cambridge

Five experiments are described in which subjects were asked to search visual displays and report the appearance, at any location on the display, of a spot of light which persisted for one second. On displays with and without a rotating radial line it was found that subjects detect fewer such spots of light near the periphery of the display. By modifying a radar-like display in order to facilitate a changed pattern of visual search, it was demonstrated that more peripheral spots could be detected.

I. Introduction

Remarkably little is known concerning techniques for scanning visual displays on which a signal may appear at any point with equal probability. Visual search is of extreme importance in the case of radar displays when employed to give warning of approaching ships or aircraft. The type of display chosen for study here was a P.P.I. (Plan Position Indicator) which is not unlike a circular TV screen, though the 'picture' of the area covered is 'painted on' by a revolving radial line of light. The 'picture' fades in a few seconds, only to be 'repainted' on successive revolutions of the radial line. A target is represented by a small spot of light (a pip) brighter than the background 'picture'. Range of a target so displayed is represented by the distance of the pip from the centre of the P.P.I., a target at maximum radar range being represented by a pip at display periphery. Bearing to the target is represented by compass marks around the periphery. Radar operators must continually search for pips. That search behaviour of operators is far from optimum is indicated by the fact that approaching targets are often not detected until their associated pips are well in towards mid-radius of the display. Such 'peripheral blindness' results in decreased warning when targets are approaching. The present studies were undertaken to determine (1) the nature of visual search behaviour in a radar-like situation, and (2) ways of biasing search behaviour in an attempt to overcome 'peripheral blindness'.

II. Method

There is, to date, no reliable way of recording eye movements over prolonged periods while subjects remain physically free, i.e. physically independent of apparatus. However, pertinent data can be obtained in a free setting by injecting signals onto a display at desired locations and having subjects respond whenever a signal is seen. This was the method employed. Subjects sat in a darkened cubicle, 34 by 40 in., facing an opal glass screen set into the narrow side of the cubicle. A steady masking noise, (58 db.), and air circulation were provided by an exhaust fan. From outside the room a circle of light 12 in. in diameter was projected onto the screen. In all experiments a pip (4 mm. diameter) appeared at predetermined intervals for 1 sec. at any one of the 16 locations shown in Fig. 1. Luminance of the projected circle was 12 foot-lamberts, while that of the pips was 14·5 foot-lamberts. When seated upright in their fixed chair, subjects' viewing distance was 18 in. Display conditions were changed slightly from one experiment to the other, though these were the ones employed in Exp. 1.

In each experiment sixteen subjects were employed, each for 64 min. A latin-square design permitted each subject to be presented with a pip at each of the sixteen locations in random order of succession

during the first 32 min. The second 32 min. period was a replication of the first. (In all, eighty R.N. ratings, sixteen in each experiment, were employed as subjects. None of them had radar experience with a conventional P.P.I. All were given sufficient practice to ensure that they thoroughly understood the nature of their task.)

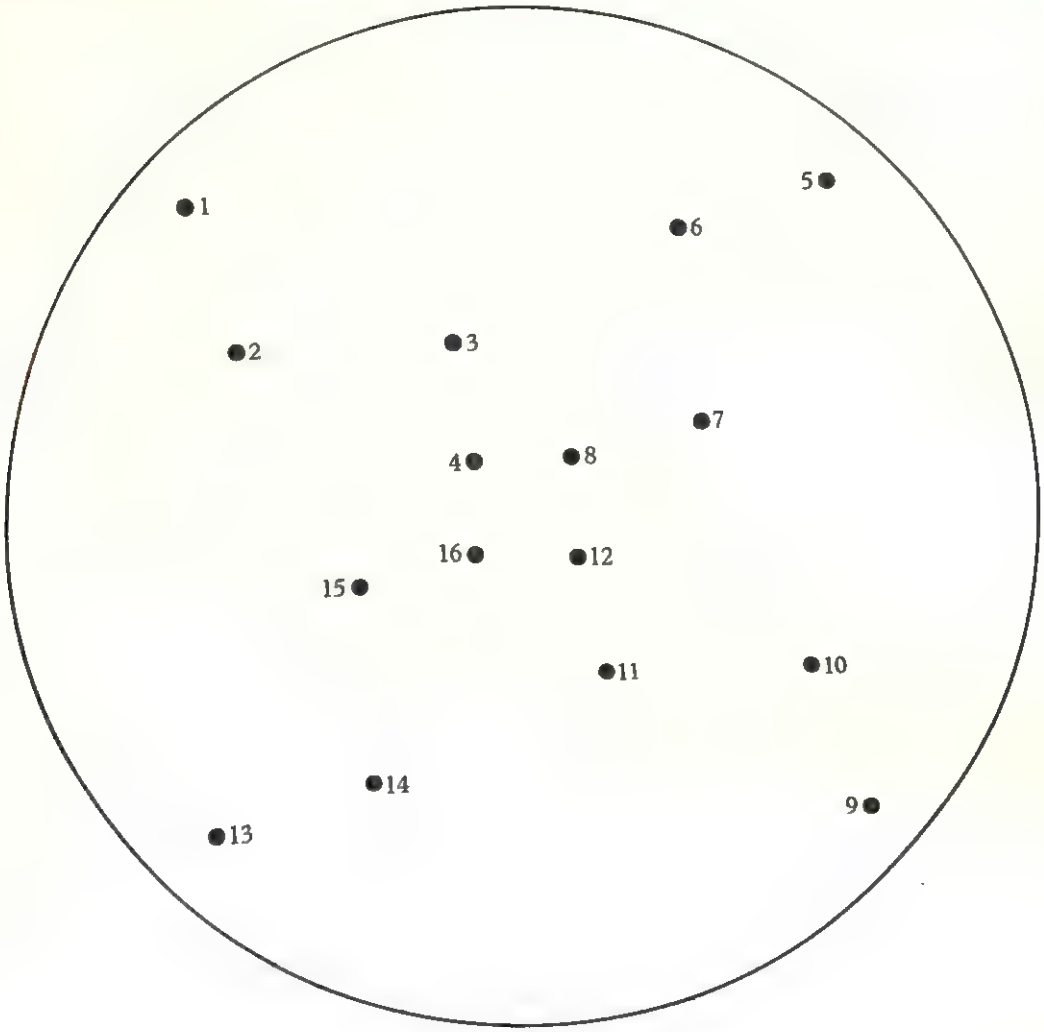


Fig. 1. Showing the sixteen locations at which the pip appeared.

Though pips appeared on an average of once every 2 min., the actual intervals between appearances ranged randomly between 36 and 196 sec., as follows: 1:58 (1 min. 58 sec. from the beginning), 3:34, 5:42, 7:04, 9:50, 11:38, 12:40, 15:36, 17:00, 18:38, 21:28, 22:44, 26:00, 26:36, 29:52, 31:30, and repeat, i.e. 33:58, etc. The subject's task was simply to search the display for the appearance of a pip and to press a switch (giving response time on a clock) whenever one was seen. They were not informed of how many pips would be presented, or where they would appear, though they were informed that the task lasted 64 min. Watches were removed.

Because the data from the five experiments reported here were later pooled for analysis purposes, the experiments are individually described below with a brief summary of results, followed by a section giving the pooled results.

III. Experimental results

Experiment 1. *Blank display (no sweep-line)*

Exp. 1 was designed to determine the distribution of visual attention to the display in the absence of any form of rotating radial sweep-line—see inserted figure for Exp. 1 in Fig. 2. Sixteen naval ratings with no experience in radar operation served as subjects. The data from Exp. 1 (see Fig. 2) indicate that

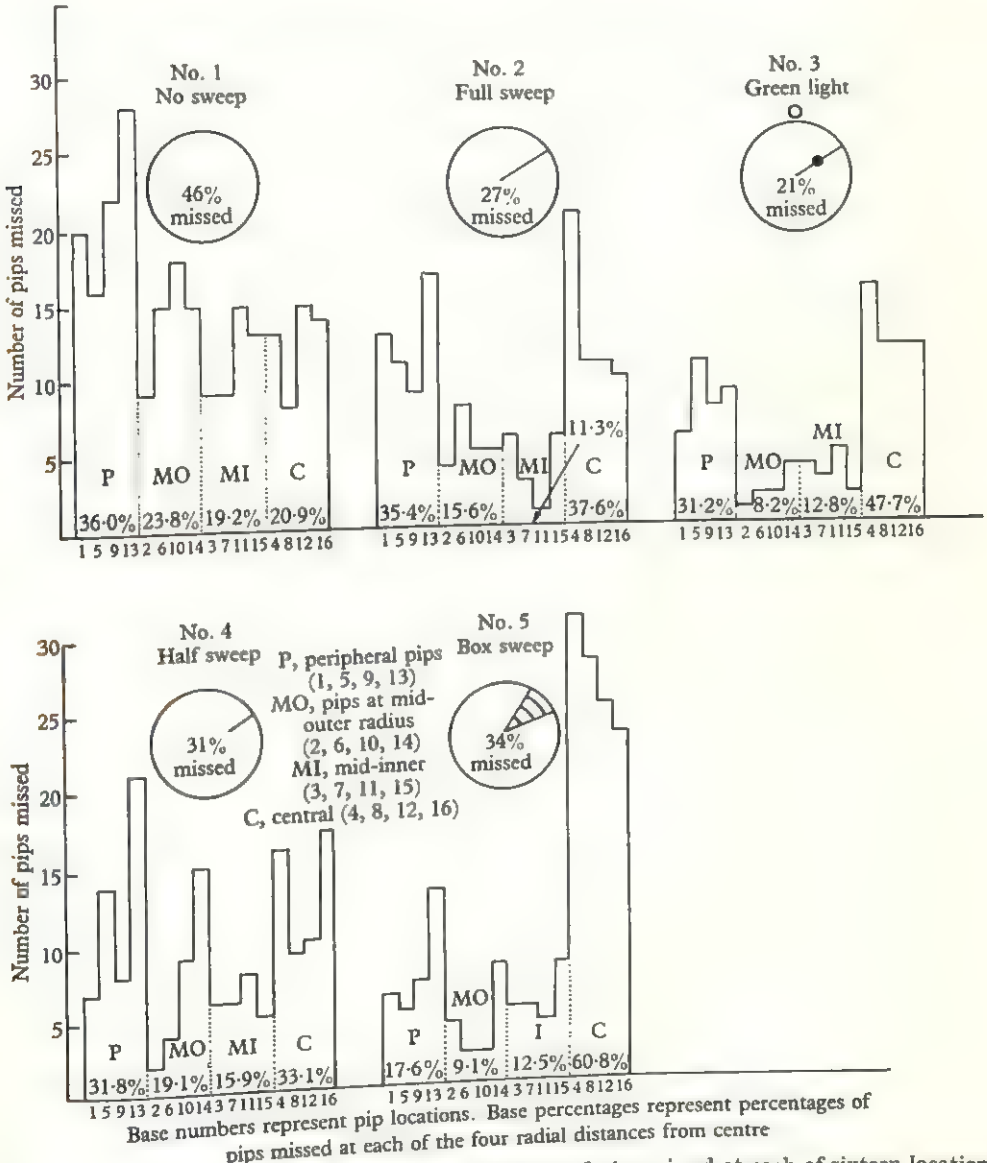


Fig. 2. Showing, for each of five experiments, the number of pips missed at each of sixteen locations. The base percentages refer to the percentage of pips missed at each of four radial distances from the centre of the display.

46% of the pips were missed, and the histogram shows the manner in which these misses were distributed for each of the sixteen locations. The greater proportion of failures to detect pips, 36%, were those located near the periphery (P). The remainder of those not detected were quite evenly distributed over the middle-outer (MO), middle-inner (MI), and central (C) regions. (Unfortunately, we know of no statistical

test which can be used to determine whether or not there are significant differences between the number of failures in detection in the different regions. However, an analysis was performed on response times, as will be shown.) Apparently there is some 'peripheral blindness' when viewing even a blank display.

Experiment 2. *Conventional* P.P.I. *display* (*full sweep-line*)

In Exp. 2 a radial line (sweep-line) was introduced which revolved at 6 r.p.m. Pips were 'painted on' now by the sweep-line—see inserted figure in Fig. 2. Subjects therefore had a much smaller area to search, i.e. immediately behind the line, and so it was anticipated that fewer pips would remain undetected. The data are shown in Fig. 2; 27% of the pips were not detected, but these were distributed differently from those in Exp. 1. As shown, there was a marked 'peripheral blindness' (35·4% of those undetected), but also a marked 'central blindness'—37·6%. Pips at approximately half radius were detected about twice as often as peripheral or central pips.

Experiment 3. *Green light*

From the data in Exp. 2 it would appear that subjects tend to search the area behind the mid-radius of the sweep-line with relatively occasional deviations to the extremes of the line from the preferred mid-areas. Exps. 3–5 introduced modifications to the display which were intended to bias subjects' search technique in favour of the peripheral regions of the display. A small dot was added to the sweep-line at exactly half radius. In addition, a small green light was positioned an inch above the display; see inserted figure in Fig. 2. The green light came on for randomly determined intervals of 2, 3, or 4 min., and was then off for any one of these periods. During the 64 min. it was on for a total of 32 min.

Subjects were instructed that when the light was off they were to search behind all portions of the sweep-line for pips (as in Exp. 2), but when it came on they were to search only from the spot to the periphery. Thus about three-quarters of the 64 min. period would be spent in searching the outer half of the display. The data (Fig. 2) show that only 21% of the pips were not detected and that visual attention was indeed biased in the required direction, i.e. 31·2% of those not detected were peripherally situated versus 47·7% centrally located.

Experiment 4. *Half-sweep*

Could greater bias be obtained than that achieved in Exp. 3? Since subjects display a preference for the region behind the mid-radius, then if only the outer half of the sweep-line was displayed subjects should continue to detect twice as many pips behind the mid-region of the half sweep-line—in this case peripheral pips. The data (Fig. 2) reveal that 31% of all pips escaped detection, 31·8% of these being peripheral compared with 33·1% in the central regions. And the misses near half radius were more numerous than those obtained in Exps. 2 and 3. These data were interpreted to mean that the conclusion from Exp. 2 was in error; subjects do not tend to search the area behind the mid-radius of the sweep-line with occasional deviations to the extremes.

Experiment 5. *Box-sweep*

The results of Exp. 4 led to the formulation of the following hypothesis. Subjects continuously search back and forth behind the line, having no particular preference for any area. With reference to Fig. 3A, let it be assumed that subjects employ four visual fixations (1, 2, 3 and 4) to 'cover' the area immediately behind the sweep-line once. Then, on returning, they use their 5th, 6th, and 7th fixations, at which point they reverse, using their 8th, 9th, and 10th fixations to reach the periphery again. It is readily apparent that over any prolonged period of time almost twice as many 'looks' are given to the mid-regions as to the extremes. The expected result would be half as many pips missed in these mid-regions, as found in Exps. 2–4. In order to distribute the 'looks' evenly it would be necessary to start the second and subsequent series of four 'looks' from the *same end* of the line each time—see Fig. 3B.

A box-sweep formed of two sweep-lines and three arc lines was employed as shown in Fig. 2. The second ('ghost') sweep-line was 40° behind the leading line. Subjects were instructed as follows: 'Search in the big box, the next, the next, and then the smallest. Then start again from the big box, and so on.' The data are shown in Fig. 2. Although a total of 34% of the pips were not detected, only 17·6% of them were peripheral compared with 60·8% in the central region. The remarkably large percentage of missed detections near centre remains unexplained.

In summary we can say that the green light technique (Exp. 3) resulted in fewest missed detections while the box-sweep resulted in greatest peripheral bias. Considering the *number* of peripheral misses only, there were thirty-four in Exp. 3 and thirty-one in Exp. 5, a non-significant difference.

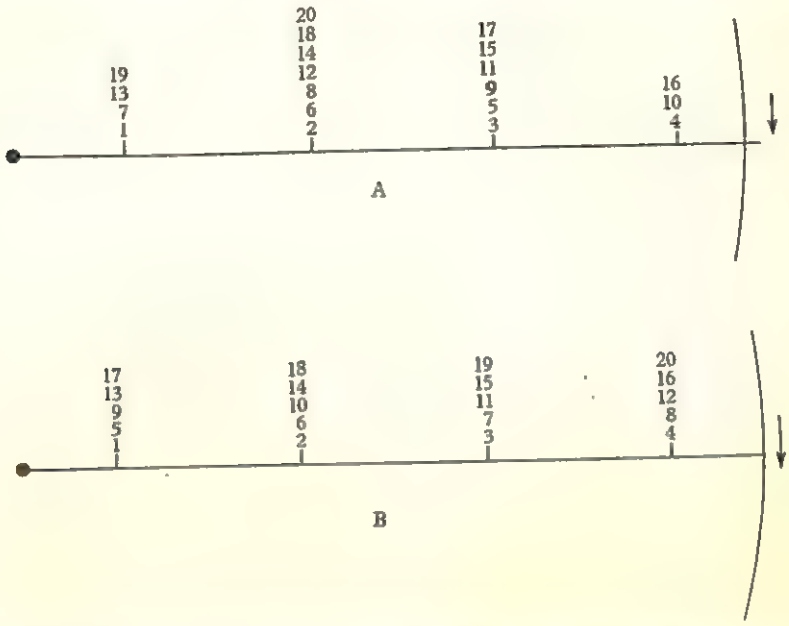


Fig. 3. (A) shows the hypothetical manner in which successive visual fixations are distributed in visually searching back and forth behind the sweep-line. (B) shows the distribution of fixations when each search is started from the same end of the sweep-line.

IV. Pooled results

The brief results given above are in terms of percentages of pips not detected. The major analysis was performed on response times. Response times for the five experiments were first transformed to log values. The greatest response time was 2·95 sec. and so the 'no-response' scores (pips not detected) were credited with a response time of 3·00 sec. Variances of each replication of each experiment were then analysed. The levels of significance obtained are summarized in Table 1.

Table 1. *Showing the levels of significance obtained from the ten variance analyses of the first and second replications (R1 and R2) for each of the five experiments*

Source of variation	No sweep		Full-sweep		Green light		Half-sweep		Box-sweep	
	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
Order	0·05	N.S.	N.S.	N.S.	N.S.	N.S.	0·05	N.S.	N.S.	N.S.
Subjects	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001
Pip radii	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001	0·001
Pip azimuth	0·01	0·05	N.S.	0·05	N.S.	N.S.	N.S.	0·05	0·001	N.S.
Radii/azimuth interaction	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	0·05	N.S.	0·01	N.S.

It is apparent from Table 1 that range/azimuth interactions occur only twice, each during a first replication. A separate analysis revealed that for the half-sweep the interaction was due to a disproportionate number of missed detections near centre in the

bottom quadrant, while in the box-sweep there were disproportionately more misses near centre in both the bottom quadrants.

The major sources of variation are subjects and pip radii. Azimuth position of pips is significant in only half the replications: in the two cases where azimuth is highly significant (no-sweep, R1, and box-sweep, R1) the level of significance is less in the second replications (0·05 for no-sweep), or nil in the case of the box-sweep. A separate analysis, not shown here, was undertaken to determine the nature of these azimuth differences. It was found that in the case of no-sweep the significantly longer response times were for the bottom left and bottom right quadrants in both replications. For the full-sweep the significantly longer times were in the top-right quadrant, while for the half-sweep and box-sweep the significantly longer times were for the bottom-right, and for the bottom-right and bottom-left quadrants, respectively.

In only two of the replications is order significant, both at the 0·05 level. In both cases this level of significance was obtained in the first half of the task, becoming nil in the second.

Before pooling the data from the five experiments Bartlett's test of homogeneity of residual variance (Edwards, 1950) was applied to the residual variances obtained by the analyses summarized in Table 1. The differences were found to be not significant, i.e. gave no evidence against the hypothesis of random sampling from a population having a common variance. Accordingly, the data were pooled, resulting in the analysis given in Table 2.

Table 2. *Variance analysis of the data from the five experiments combined*

Source	D.F.	S.S.	M.S.	*F*	*P*
Experiments (E)	4	9·773	2·4433	8·339	0·001
T × L (between)	15	6·196	0·4131	1·409	N.S.
Error (between)	60	17·577	0·2930	—	—
Subjects (S)	79	33·546	0·4246	17·052	0·001
Replications (R)	1	0·396	0·3960	3·535	N.S.
Trials (T)	15	0·675	0·0450	1·034	N.S.
Locations (L)	15	12·169	0·8113	5·879	0·001
T × L (within)	210	6·621	0·0315	1·265	0·05
T × E	60	2·609	0·0435	1·747	0·01
L × E	60	8·278	0·1380	5·542	0·001
R × E	4	0·448	0·1120	4·498	0·01
R × L	15	0·434	0·0289	1·161	N.S.
R × T	15	0·442	0·0295	1·185	N.S.
R × L × T	225	5·813	0·0258	1·036	N.S.
R × E × T	60	1·295	0·0216	0·867	N.S.
R × L × E	60	1·429	0·0238	0·955	N.S.
E × L × T (within)	840	26·613	0·0317	1·273	0·01
Error (within)	900	22·452	0·0249	—	—
Total	2559	123·220	—	—	—
Azimuth (A)	3	0·638	0·2127	1·541	N.S.
Radius (Rad)	3	10·015	3·3383	24·190	0·001
A × Rad	9	1·516	0·1684	1·220	N.S.

Starting with the larger interaction terms, it is apparent from Table 2 that E × L × T is highly significant. We interpret this to mean that there is indeed a systematic E × L × T interaction, or an S × L × T interaction, or both. Despite this significance it is apparent from the *F*-ratio that neither effect is large in comparison to the experimental error. The three remaining second-order interactions are not significant, showing that R × L × E, R × E × T, R × E × L, and the subject interactions S × R × L and S × R × T are not significant. The non-significance of the R × T interaction supports the view that the S × R × T

interaction is not significant, while that of the R × L supports the view that S × R × L is not significant.

The R × E interaction, being significant, suggests that either there is a systematic R × E interaction, or an S × R interaction, or both. Unfortunately, we cannot distinguish between these possibilities. Again, the E × L interaction being significant, could also imply an S × L interaction.

The significance of the T × E interaction could be due, too, to an S × T interaction, while that of the T × L (0·05 level) could be due to an S × T × L interaction.

It was noted above (L × E interaction) that an S × L interaction may exist. Consequently we have tested locations (L) against L × E, yielding an F-ratio of 5·879. Evidently the effect of pip location is very marked. This location effect has been further analysed into azimuth, radius, and A × Rad interaction, as shown in the lower portion of Table 2. From this analysis it is readily apparent that the significant effect of location is almost totally due to radius.

With reference to the significance of T (Trials), it was noted that an S × T interaction might exist. Hence, we have tested T against T × E. The resulting F-ratio of 1·034 is not significant. And for the reason that an S × R interaction might exist we have tested R against R × E. The obtained ratio of 3·535, being non significant, indicates no effect of replications. The effects of subjects and of experiments are both highly significant.

In summary, then, it can be said that when response times for the five experiments are analysed together, experiments, subjects, and locations are highly significant factors, with one aspect of location, pip radius, contributing more to the variance than any other factor.

V. Discussion

Distribution of attention

While a few studies have been made of visual scan techniques when regarding displays having certain static attention-getting features, e.g. Brandt's (1945) data using pictures, and the study by Fitts, Jones & Milton (1950) using aircraft instrument displays, the present studies are the only ones known to the writer in which the appearance of a signal at any point on the display is equally probable. Here we have subjects in a free setting, searching a display on which a brief signal *might* occur at any location.

The evidence from Exp. 1 is that when viewing a blank display subjects tend to disregard the periphery, while distributing their attention fairly evenly over the remainder. When 1 or 2 min. have elapsed since the preceding signal, Broadbent (1950) found that when watching a horizontal line of twenty dials, 'the centre of the display is more closely observed than the ends'. Broadbent's subjects reversed this tendency when signals were more than 5 min. apart. Broadbent's signals, however, remained on until detected. In another dial-watching experiment Lincoln & Averbach (1956), too, found that signals on centrally located dials were detected more often than those peripherally located, when the dials were tachistoscopically exposed.

These two experiments offer confirmatory evidence of 'peripheral blindness' when searching for a possible event. Does the same phenomenon obtain when searching for one among a number of events simultaneously displayed; for instance, when searching for one of a large number of clearly legible digits? This problem was investigated by Green, McGill & Jenkins (1953), who concluded that, 'certain numbers in a configuration are

easier to find than others', but they were unable to analyse for location. Accordingly, we undertook a sixth experiment, as follows.

A black circle having an area of 100 sq. in. was drawn on white cardboard. In the circle were printed 100 two-figure digits, 00–99, one pair of digits being randomly placed within each square inch of the display. All digits were oriented in the same direction. Sixteen pairs of digits were selected having the same relative locations as the pips in Exps. 1–5 (Fig. 1). Thirty subjects were individually tested to determine the time taken to locate these numbers, which were given one at a time. A variance analysis of the log values of the obtained times showed that subjects took significantly longer (0·05 level) to find the four peripherally located numbers than the remaining twelve.

Hence, it is argued that the phenomenon of 'peripheral blindness' characterizes visual search techniques when searching for an event or when decision taking, i.e. selecting one from a number of events according to some given rule.

It was shown in Exp. 2 that when a rotating sweep-line was introduced attention was distributed along the area behind the line in a manner reminiscent of a normal distribution. (It must be noted that the speed of rotation of the line is an important factor: at speeds too great for the eye to 'follow' the distribution would doubtless have been different.) Here, as indicated, the search area was much smaller than in Exp. 1. Buckley, Hanes & Deese (1953) have shown that probability of pip detection increases with decrease in area of search, hence the greater number of detections in Exp. 2. Because the area of pip appearance (and of visual search) was immediately behind the line it is reasonable to conclude that 'peripheral blindness' characterizes this smaller area too, i.e. 'blindness' at the extremes of the line rather than of the whole display. It appears that the reason radar operators often fail to detect targets at extreme radar range is simply that they do not distribute sufficient visual attention in peripheral display regions.

Smith & Boyes (1956) have since confirmed the phenomenon of 'peripheral blindness' on an actual radar display. They reproduced the display conditions of Exp. 2, using a 12 in. cathode-ray tube and a sweep-line revolving at 6 r.p.m. Pip thresholds (*detectability* thresholds) were determined by increasing the signal voltage 1 db. on each rotation of the sweep-line until the pip was detected. In addition they investigated three search areas—full display, half display, and quarter display, That is, subjects were informed that the pip would appear anywhere on the scope, in some particular half, or in some particular quarter.

The results are shown in Fig. 4. It is quite apparent that pip intensity must be markedly greater if it is to be detected near centre or periphery than if near half-radius. (In addition it should be noted that detection was better—threshold signal intensity lower—when only half or quarter of the display was being searched.)

The problem remained of determining whether this distribution of attention was due to subjective behaviour or to some characteristic of the cathode-ray tube. Accordingly, *visibility* thresholds were obtained using the same cathode-ray tube, i.e. subjects were informed as to precisely where the pips would appear and threshold voltages were again determined. The result is shown in the top curve of Fig. 4. The slight drop in threshold voltage near centre remains unexplained. Of greatest importance, however, is the fact that *visibility* thresholds near the periphery were just as low as those near mid-radius. We are therefore led to conclude that the results of Exp. 2 can indeed be extrapolated to the real radar setting.

Exps. 3 and 5 demonstrated that the distribution of attention to such displays can indeed be biased in a desired direction—that 'peripheral blindness' can be significantly decreased by relatively slight modifications of the display. It should be noted here that in instructing subjects how to search in the green-light or box-sweep experiments we have no objective evidence that they did in fact do as instructed, though all subjects, when later questioned, maintained that they did. Furthermore, it is evident from an analysis

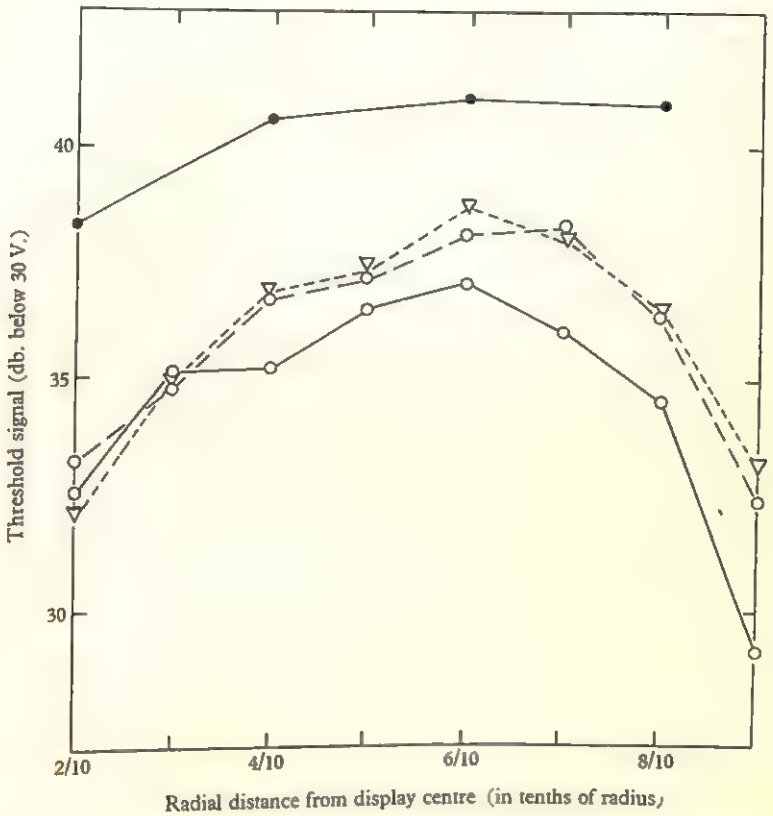


Fig. 4. Showing detectability thresholds for three areas of search, and visibility thresholds (top curve), as a function of radial distance of a pip from the centre of a radar display. The ordinate values are in decibels attenuation of a reference voltage: the greater the value the lower the threshold. Visibility, ●——●. Detectability: O——O, full scope; O— —O, half scope; ▽ - - - ▽, quarter scope.

not given here that they maintained their search techniques, as instructed, throughout the hour—the histograms in Fig. 2 are representative of both replications in each of these two experiments. An unexplained feature of Exp. 5 (box-sweep) is the large number of misses in the central region. One possible reason is that because this inner box is the smallest it was 'searched' by only a single visual fixation. Another suggestion is that triangular displays are given less visual coverage than are those with four or more sides.

The data from Exp. 5 support our hypothesis. During a radar-like task subjects search back and forth in the area where pips are likely to appear. This method of search has been termed 'search by parallel sweeps' by Koopman (1956).

VI. Conclusion

When viewing a blank display on which a signal might occur at any location, subjects detect fewer signals occurring near the periphery of the display. Such 'peripheral blindness' obtains, too, when searching for one of a display of digits. In a radar-like task where signals might occur behind a rotating radial line, 'peripheral blindness' occurs near the two extremes of the line: signals near mid-radius are detected about twice as often. Visual attention can be biased toward the display periphery by changing the pattern of visual search.

This study was undertaken, while the author was on a tour of duty with the Medical Research Council, from the Defence Research Medical Laboratories, Defence Research Board of Canada, in Toronto. The work was done under the general direction of Dr N. H. Mackworth, and generous statistical advice was given by Mr M. Stone.

References

Brandt, H. F. (1945). *The Psychology of Seeing.* New York: Philosophical Library.

Broadbent, D. E. (1950). The twenty dials tests under quiet conditions. A.P.U. Report No. 130/50.

Buckley, B. B., Hanes, R. M. & Deese, J. (1953). Search area and target detectability on a PPI cathode-ray tube. Wright Air Development Center, Tech. Report No. 52–303.

Edwards, A. L. (1950). *Experimental Design in Psychological Research.* New York: Rinehart.

Fitts, P. M., Jones, R. E. & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument landing approaches. *Aero. Eng. Rev.* **9**, 1–16.

Green, B. F., McGill, W. J. & Jenkins, H. M. (1953). The time required to search for numbers on large visual displays. M.I.T., Lincoln Laboratory. Tech. Report No. 36.

Koopman, B. O. (1956). The theory of search. II. Target detection. *Operat. Res.* **4**, 503–31.

Lincoln, R. S. & Averbach, E. (1956). Spatial factors in check reading of dial groups. *J. Appl. Psychol.* **40**, 105–8.

Smith, A. A. & Boyes, G. E. (1956). Radar target detectability as a function of search area and range. Defence Research Board of Canada, D.R.M.L. Report No. 163–8.

(*Manuscript received* 9 *July* 1957)

EXPLORATORY BEHAVIOUR

By S. A. BARNETT

Department of Zoology, Glasgow University

A critical analysis is presented of some features of exploratory behaviour, especially in wild and tame rats but also in other species. Such behaviour is considered in the light of certain general concepts, especially 'appetitive behaviour', 'reactive inhibition', latent or exploratory learning, and motivation. The avoidance of new things, displayed by wild rats, is also discussed.

I. Introduction

That laboratory rats are highly exploratory and 'inquisitive' has long been known, though the importance of this behaviour has only gradually been realized. Small (1899) described the restlessness of infant rats as 'premonitions of curiosity', and remarked that fully developed exploratory behaviour, evident from the age of 3 weeks, is not merely a reflexion of hunger. In the modern period of behaviour study, among the first important observations were those of Blodgett (1929): this work was the first of a series of studies, by various authors, on 'latent learning', demonstrating that the exploratory movements of rats are often accompanied by topographical learning which is independent of evident reward. Reviews of this subject have been published by Majorana (1950), Munn (1950, pp. 315 *et seq.*), Thistlethwaite (1951) and Thorpe (1956, pp. 90 *et seq.*). Apart from workers concerned with latent learning, many psychologists have remarked in passing on the exploratory propensities of white rats (e.g. Buytendijk, 1931). Occasionally, this type of behaviour has obtruded itself where something else was sought: Locke (1936), for instance, looking for evidence of a 'social drive', found instead marked indications of an 'exploratory drive'. Since 1945 the study of exploration has been given more attention, and this recent work provides much of the substance of the discussion below. Independently of psychologists working with laboratory rats, zoologists have begun to study wild rats, both in the field and in cages or enclosures. While psychologists have, on the whole, used rats as 'preparations' showing special phenomena such as latent learning or spontaneous alternation, zoologists have been concerned with *Rattus norvegicus* Berkenhout (the species to which tame rats also belong) as an animal which maintains itself in certain environments (e.g. Chitty & Shorten, 1946; Calhoun, 1949; Barnett & Spencer, 1951; Chitty, 1954). Further, by using wild rats, they have been able to study a greater range of behaviour than is shown by tame, laboratory-bred animals: in particular, wild rats display a type of avoidance behaviour not observed in tame ones (Chitty & Shorten, 1946; Shorten, 1954; Barnett, 1956, 1958*b*). The object of this paper is to present a critical analysis of some features of exploratory behaviour, especially in wild and tame rats but also in other species; and to relate this to certain general concepts, especially 'appetitive behaviour', 'reactive inhibition', latent or exploratory learning and motivation; the avoidance of new things, displayed by wild rats, is also discussed. Exploratory behaviour is of great importance in many, perhaps most, animal species. It has been studied in a cockroach, *Blattella germanica* (Darchen, 1952, 1955), the hive bee, *Apis mellifera*, (Lindauer, 1953), the rhesus monkey, *Macacca mulatta* (Butler, 1953, 1954; Butler &

Alexander, 1955; Harlow, Blazek & McClearn, 1956), and the chimpanzee, *Pan troglodytes* (Welker, 1956*a*, *b*). But the most detailed and extensive studies have been those on rats, and it is these that will be principally drawn upon in this review.

II. 'Appetitive' and exploratory behaviour

(1) *Definition of 'appetitive'*

Many animals perform variable (unpredictable), sometimes apparently random, movements before they reach a situation in which they can carry out a stereotyped act such as eating. The name given to such movements is *appetitive behaviour*; a formal definition of this term in common use is: 'the variable introductory phase of an instinctive behaviour pattern or sequence' (Thorpe, 1951). This type of behaviour has been very fully discussed by Thorpe (1956, pp. 36 *et seq.*); as he points out, the activities so defined are exceedingly diverse and, in their relations to the various stereotyped acts, very complex. A simple example is the quartering of its territory by a predatory bird, before a victim is seen, pursued and captured; but not all appetitive behaviour is on this scale, since the movements, within a small compass, of, say, a bird shifting about before settling in its nest, also come within this class of actions. In the discussion below only the larger-scale movements about the animal's living space are considered, but this is for convenience only.

The definition given above reserves the term 'appetitive behaviour' for movements in an *instinctive* sequence. This would presumably exclude random movements preceding any activity in the development of which learning had taken a substantial part. For instance, Seitz (1940) brought evidence that in Cichlid fish of the genus *Tilapia* the males do not have an inborn ability to recognize conspecific females, but that they learn to do so during early life. If this is so, then courtship by the males of this genus depends on learning, and it might be questioned whether this particular behaviour should be called 'instinctive'. Admittedly, in this example the motor performance of the male is stereotyped for a given species; but in other instances it may be that even stereotyped action patterns, such as bird song (Thorpe, 1954*b*), or reproductive behaviour in rats (Beach, 1951), can be performed only if the animal has had the opportunity to learn them. Another way of putting this point is to say that definitions of the term 'instinctive behaviour' beg the question of just how the behaviour referred to develops in the ontogeny of the individual (cf. Lehrman, 1953): consequently, it is sometimes difficult to decide whether a specific behaviour pattern comes within the scope of the definition. As far as defining 'appetitive behaviour' is concerned, the obvious way out of this difficulty is to omit the word 'instinctive', and so to use 'appetitive behaviour' to refer to variable movements preliminary to any stereotyped activity, regardless of the extent to which 'learning' has played a part in its development.

However, even when the definition is shorn of this particular ambiguity, a serious difficulty remains in describing variable behaviour as it is actually observed. This is because behaviour which, while it is going on, is indistinguishable from appetitive behaviour as defined, may end 'spontaneously': that is, as a result neither of general fatigue nor of the performance of a consummatory act (Hinde, 1953; van Iersel, 1953; Kortlandt, 1955, pp. 209, 213; Thorpe, 1956, p. 41). In Kortlandt's picturesque phraseology: 'it looks as if Nature mistrusted appetitive behaviour and therefore has added to it some interim reward'. In other words, this variable activity is sometimes an end in itself, and

the distinction between appetitive and consummatory behaviour cannot be sharply drawn.

Nevertheless, even when random or wandering movements do not end in a specific consummatory act, they can often be shown, or at least inferred, to be instigated by an internal state representing a biological need; in view of this, and of the associations and etymology of the word 'appetitive', it might be thought appropriate to use the term to name those variable movements of an animal which are evoked by an internal state corresponding to a specific primary need (including the need to reproduce). The behaviour would then be defined in terms of physiological state rather than end-result.

The termination of behaviour set in motion by a specific internal state has in the past been ascribed to the *performance* of a stereotyped *activity*, such as drinking, ejaculation or nest-building; but it is now widely conceded that it is not necessarily the performance of an act which is significant, but achieving a particular input to the central nervous system. On this view, the termination of a behaviour sequence is due to the attainment of a consummatory *state*, or afferent pattern (exteroceptive or interoceptive or both) falling on the central nervous system. A consummatory state may result from performing some final act which satisfies a biological need; or it may result merely from performing a preliminary to such an act; or it may arise from the attainment simply of a sensory condition, such as a particular skin temperature in a mammal (Hardy, 1954) or a certain arrangement and number of eggs in contact with a bird's ventral surface (e.g. Moynihan, 1953). The concept of consummatory state has been discussed in some form by Bastock, Morris & Moynihan (1953), by Thorpe (1954*a*; 1956, p. 105), by Kortlandt (1955) and by Barnett (1955). (It is further considered below in VII, 2.) The physiology of such a state must, in the present state of knowledge, often be hypothetical, but there is nothing new in that.

(2) *The variable movements of rats*

Further analysis of 'appetitive' behaviour may be based on an examination of the variable movements of rats about their living space. In wild rats, a generalized movement about a wide area is most readily evoked in two types of situation: first, if a new area becomes accessible, for instance by the removal of a barrier (e.g. Barnett & Spencer, 1951); second, if rats are introduced into an unfamiliar cage or enclosure (Barnett, 1958*a*). When they are in familiar surroundings the behaviour still occurs, though at a lower intensity, and is then most marked when the animals are satiated (Barnett, 1956). As will be seen below (VI, 3), latent learning, or exploratory learning (Thorpe, 1944), is similarly best demonstrated in albino rats when they are not hungry or thirsty or otherwise deprived. If the 'official' definition of 'appetitive behaviour' is adhered to, it is clear that exploratory behaviour released by satiation cannot properly be called appetitive. If it were desired to include 'unmotivated' exploratory behaviour under the heading of appetitive activity, then the definition of 'appetitive' would have to be changed, even if the difficulties outlined in the preceding section were ignored.

It may still be asked to what extent rats display appetitive behaviour *sensu stricto*. That they do so in small-scale movements, for instance in the preliminaries to coitus, is not in doubt, but the answer is less clear when wider-ranging movements are considered. It has long been known that the *activity* of rats is liable to increase in certain physiological states, for instance during oestrus (Richter, 1927). The same effect has recently been demonstrated by Finger & Reid (1952) in rats deprived of food or water. This might seem

to be a measure of appetitive behaviour. But the activity observed, running in a wheel, is not exploratory, but has rather the character of a compulsive motor performance: it may be substitutive, and so analogous to displacement behaviour (Eayrs, 1954). This requires further investigation. In any case, the motion of a wheel tends to stimulate further movement by the rat, once running has begun, and so exaggerates any existing tendency to be active.

It has been doubted whether deprivation by itself increases activity, but this doubt has now been resolved. Jerome, Moody, Connor & Fernandez (1957) have studied the movements of albino rats in a cage with two compartments connected by three doors. To cross from one compartment to another, a rat had to find the one door that was unlocked. The number of crossings was increased by bright light (which was avoided) and by hunger; but even in the absence of any such motivation, many crossings were made. In these experiments, then, an 'appetitive' effect was superimposed on a less labile tendency to explore regardless of specific need.

Campbell & Sheffield (1953) had earlier drawn attention to the importance of environmental change in influencing activity (see also III, 2 below). They studied the 'random activity' of rats in a constant environment, and in their experiments food deprivation alone was followed by little increase. If, however, the deprivation was accompanied by environmental change, there was a great increase. They concluded that 'hunger drive' brings about a lowered threshold of response to change, but is not in itself an internal stimulus to activity. Adlerstein & Fehrer (1955) came to the same conclusion on the basis of experiments with rats in a complex maze. Still more recent work has confirmed that these workers are right in attaching importance to lowered threshold of response to environmental change, but has combined a demonstration of this 'arousal' effect with further proof of the influence of physiological need. Hall (1956) tested rats in activity wheels (*a*) in the dark, and (*b*) while subjected to external stimulation from flashing lights and noises; some of the rats were hungry, some sated. In these conditions, rats did show increased activity resulting from external stimulation, but a still greater increase attributable to hunger. Strong (1957) has confirmed once again the increased activity due to hunger; he has further shown that the activity affected is of the gross locomotor type and not fine, non-locomotor movements. This work emphasizes the importance of the particular method used in measuring 'activity'.

Fehrer (1956), perhaps unusually for a psychologist, has drawn attention to the biological function of exploration. She suggests that, to test the effect of internal state on exploration, it is better to give rats the opportunity to enter a new place from an unfamiliar one, than to transfer them from a cage to a maze. In her experiments, hungry rats showed a greater readiness to leave a familiar place than did sated rats.

Evidently activity in rats is evoked by two types of stimulus situation. First, environmental change, influencing the exteroceptors, increases exploration. Secondly, deprivation (*a*) lowers the threshold of response to external change, (*b*) in itself may increase, not only gross locomotor activity in general, but also exploration proper. Some theoretical implications of this are discussed later (VII, 2).

(3) *The classification of variable movements*

From the preceding survey it follows that there are two ways or dimensions in which the variable movements of an animal might be classified. First, appetitive behaviour

might be distinguished from exploration: in this case the distinction could be based on (hypothetical) internal states, since the term 'appetitive' would be applied only to behaviour 'motivated' by physiological processes related to certain specified biological needs. We have already seen what difficulties arise if this sort of classification is used.

Secondly, there is a possible distinction between (*a*) the large-scale movements of an animal about its living space, and (*b*) the minor variable movements which immediately precede some specific act or the attainment of a specific state. For instance, a female rat with young will at intervals *explore* the space constituted by her home range; she will also, when necessary, search around a strayed nestling and retrieve it: the retrieving is a consummatory act, preceded by a variable phase of activity, namely, the approach to the young. There is, however, no clear boundary between the large-scale movements on the one hand and those covering only a short distance on the other.

Neither of these classifications is satisfactory, because each encounters difficulties when the attempt is made to apply it to the observed facts of behaviour. A summary statement of these facts is as follows. Animals perform variable movements covering both large distances and small. Often, these movements end when a particular nervous input (consummatory state) has been achieved (whether by performance of a consummatory act or not). Sometimes, the activity ends without the attainment of any recognizable consummatory state; but in either case it is often possible to assert, either from experimental evidence or by inference, that the activity has been set in motion by a *specific* internal process. Some of these variable movements are exploratory in the sense that they place the animal in contact with many different parts of its environment. All these diverse activities may be labelled 'appetitive'; if this is done, the main function of the label is to distinguish them from those stereotyped patterns of behaviour such as courtship ceremonies, territorial displays and so on which are, by comparison, not variable but fixed (that is, predictable in detail) within each species. As Thorpe (1956) has emphasized, it is in the variable phase of activity that learning principally takes place.

The term 'appetitive behaviour' then comes to mean *any variable behaviour*, and it is necessary to state for each instance to what 'goal object' (if any) and to what internal state (if known) it is related. Sometimes exploratory behaviour is unrelated to any 'goal object', and independent of any internal state other than that which brings about the exploratory behaviour itself. Among rats, this behaviour is observed in its purest form in the laboratory varieties, protected as they are from environmental hazards and selected for tameness (Barnett, 1958*b*). The internal state or processes involved are probably similar to those which instigate play (Lorenz, 1954; Kortlandt, 1955, pp. 225 et seq.; Thorpe, 1956, pp. 84 et seq.) Both exploration of this sort, and play, are released when primary needs are satisfied, and the animal is free to behave in an 'irresponsible' way.

III. STIMULI EVOKING EXPLORATION

(1) *Spontaneous alternation and 'reactive inhibition'*

The preceding section has been concerned with the description of exploratory and similar behaviour, and has made some reference to its internal ('physiological') causes. The next step is to examine in detail the causes of exploratory behaviour, especially the *external* stimuli which influence it. It is convenient to begin with experiments on spontaneous alteration. This phenomenon is observed in maze-running experiments with

Bureau Edn. Psy. Research
DAVID H ING COLLEGE

equally rewarded or unrewarded alternatives, and most clearly in a maze in which there is only one point at which a choice of routes has to be made. In such a T- or Y-maze, if rats are given two successive opportunities to make this choice, on the second occasion they tend to make the choice not made on the first: if a rat has turned left on the first occasion, it will probably turn right on the second. The same alternation can be observed in more complex mazes.

Montgomery (1951) pointed out the relationship of alternation with exploration: alternation, when the results in terms of reward are equivalent whichever choice of route is taken, is a special case of exploratory behaviour, in that it tends to increase the animal's range of movements. It follows, therefore, that any theoretical interpretation of spontaneous alternation is likely to have a bearing on exploratory behaviour in general.

The most important theoretical discussion until very recently was that of Solomon (1948), who suggested that alternation could be explained in terms of Hull's (1943, 1951) principle of reactive inhibition. This is formally stated as follows (Hull, 1951): 'Whenever a reaction (R) is evoked from an organism there is left an increment of primary negative drive (I_R) which inhibits to a degree according to its magnitude the reaction potential ($_SE_R$) to that response.' The empirical generalization contained in this theorem is that the performance of an act is followed by a period during which the probability of that act being repeated is reduced. This inhibitory effect dissipates gradually: as an example, Heathers (1940) showed that alternation in a maze decreased proportionally with increase in running time; for instance, with an interval between choices of 15 sec., alternation occurred in 83·5% of trials while, with an interval of 120 sec., the percentage was 65·6.

When a response is rewarded, it tends to be repeated. It was assumed by Hull that this is due to an excitatory potential which persists for much longer than the reactive inhibition. Both excitatory and inhibitory potentials, in this scheme, are theoretical constructs expressing generalizations about behaviour: they are 'intervening variables'; they may, however, be regarded as referring to processes, not yet identified, going on in the central nervous system.

Hull included in his concept of I_R the hypothesis that failure to repeat a response is due to the *performance of the response*. Montgomery (1952*a*, *b*, 1953*a*) and Glanzer (1953*a*, *b*) independently published evidence that this hypothesis is incorrect. Glanzer suggested that failure to repeat a response can appropriately be described in terms of *stimulus satiation*: with continuous exposure to an environment, that is, to the same constellation of external stimuli, the animal becomes less active in that environment. This would apply, in Glanzer's view, both to spontaneous alternation and to exploratory behaviour in general. Sutherland (1957) has developed this notion by running rats in modified T-mazes in which the goal was invisible at the choice point. In one type of maze both arms led to the same goal, in the other there were two, dissimilar goal boxes. His experiments confirmed that an important factor in spontaneous alternation is avoidance of stimuli recently experienced; further, this was the case even though the stimuli were not present to the animal at the choice point.

Walker, Dember, Earl & Karoly (1955) have made a very detailed analysis of the causes of spontaneous alternation. They used a + maze which could be rotated, and of which one arm could be blocked to give a T; they showed that in such a maze the factors responsible for alternation were partly the visible difference between the arms of the maze itself (painted black or white) and partly the visible difference between the backgrounds of the

different arms. The effect of the alternation they observed was to vary the *visual stimulus* received by the rat in successive choices, but there was no tendency to vary the *response*, that is, to alternate between turning right and turning left. In this carefully designed study the conclusion again emphasizes the importance of the stimulus pattern, and not the response pattern, in inducing alternation.

In a later study, Walker, Dember, Earl, Fawl & Karoly (1955) showed that it is possible to design an experiment in which alternation is determined partly by the response. They used a T-maze in which the entries to the two goal arms were complicated: each involved twists of the body and changes of direction, and the contortions demanded of the rat were quite different in the two arms. In these conditions, response alternation was nearly as marked as stimulus alternation. The authors suggest that an important factor is the readiness with which two alternatives can be *discriminated*; merely turning left or turning right are not very different, but one set of contortions is readily distinguished from another.

An important question is whether response alternation and stimulus alternation can both be regarded as consequences of a single fundamental process. Dember (1956) exposed rats to a T-maze in which one goal arm was painted white and one black, but he prevented them, by means of glass barriers, from entering either arm on the first run. On a second exposure, both arms were now painted the same colour (either black or white), and either could be entered. Most of the rats explored the *changed* arm. Dember therefore emphasizes the importance of *novelty*. It seems that alternation is essentially an expression of a tendency to behave so that the input to the central nervous system, whether through the exteroceptors or through proprioceptors as in the case of response alternation, is varied. It is as if the animal were continually behaving so as to avoid boredom (cf. Myers & Miller, 1954). This principle probably applies to all types of external stimulation: for example, rats given a choice of three or four foods, of which one is more acceptable than the others, nearly always regularly sample the alternatives; they tend to do this after a period of feeding on the preferred food (Barnett & Spencer, 1953*b*; Barnett, 1956). Butler (1957) finds that even 'auditory exploration' can be a factor in the behaviour of rhesus monkeys.

(2) *Attention to change*

All recent studies, not only of spontaneous alternation but of exploratory behaviour generally, reinforce the conclusion that variation in the stimulus situation is important. It might be thought, however, that general exploratory behaviour is an expression of a tendency to be active. Among those who have shown that this hypothesis is untenable are Berlyne (1950) and Montgomery (1952*b*, 1953*a*, *c*). Glanzer (1953*b*) applied his notion of 'stimulus satiation' to exploratory behaviour as well as to alternation; and Berlyne (1955) used the term 'stimulus hunger' to describe the tendency of rats (and of other species) to investigate unfamiliar objects and situations.

The experiments of Berlyne (1950) showed that rats spend more time exploring new things than things which they have already explored, and that the second time they encounter a new situation they spend less time exploring it than the first. Thompson & Solomon (1954) confirmed Berlyne's findings by showing that rats presented with a new visual pattern explored more than do rats presented with one already experienced. Montgomery (1953*a*) used mazes which were identical apart from being painted black,

grey or white, and found that the amount of exploratory behaviour shown in the second maze entered by a rat was a function of its similarity to the first. In this paper Montgomery speaks of an 'exploratory drive'. In the work so far discussed the three factors of novelty, complexity and spaciousness were not clearly distinguished. Berlyne & Slater (1957) have accordingly done experiments in Y-mazes in which these parameters could be separately studied. They found evidence that all three help to evoke exploration.

It might be thought that the effect of spaciousness reflects a tendency on the part of the rat to 'escape'—a suggestion sometimes made in discussion, even if not in the literature. But, in fact, exploration is typically followed by *return* to the nest or other central point. The notion of escape or freedom for its own sake is an anthropomorphism: as Hediger (1950) has pointed out, one must specify from what the particular individual is escaping (for example, disturbance by men or a hostile companion). Exploratory behaviour does not in itself result in a permanent exchange of one environment for another: it does so only if there is some specific motivation for such a change. When there is no such motivation, exploration merely leads the animal to experience all the accessible parts of a particular range (see also VI, 1 and VI, 3).

Further justification for expressions such as 'stimulus hunger' is found in the observations of Montgomery (1954) and of Myers & Miller (1954), who found that the opportunity to explore an area can act as a reward in learning experiments. Merely presenting novei stimuli, without increasing the space available for exploration, constitutes an effective reward (Chapman & Levy, 1957). Similarly, Dember, Earl & Paradise (1957) find that rats tend to move from conditions of less complex visual stimuli to those in which the stimuli are more complex. The strength of the tendency to explore, even when the animal is in a state of primary need, has been further demonstrated by Zimbardo & Montgomery (1957). In their experiments rats were placed in a 'checkerboard maze' of nine squares in each of which there was a cup which held food or water. In these conditions thirsty rats explored for many minutes before drinking, hungry rats for longer still before eating. These experiments recall the observation of Krechevsky (1937) that rats will choose the longer of two alternative routes to food, if the longer route is varied. This has been confirmed by Hebb & Mahut (1955) and by Havelka (1956).

There is evidence on similar lines for several other species. In rhesus monkeys it had already been shown that the opportunity either to carry out manipulations, or to 'explore' visually, can act as effective rewards in learning situations (Harlow, 1950; Harlow, Harlow & Meyer, 1950; Harlow & McClearn, 1954). Thompson & Heron (1954*a*, *b*) have shown that, in both rats and dogs, restricted opportunity for exploration in early life is followed later by greater exploratory activity than is shown by control animals. For man, Bexton, Heron & Scott (1954) have studied the effects of sensory deprivation. Reduction of external stimuli to a minimum, in comfortable conditions, led to a demand for stimulation. If this demand were not satisfied, extensive fantasies and even hallucinations were reported. At the end of the experimental period behaviour was for a time considerably disorganized. For our own species, however, it is hardly necessary to provide a formal demonstration of 'stimulus hunger' (cf. Hebb, 1955).

All these observations lead to the conclusion that exploration is a behaviour 'pattern' in its own right. Dember & Earl (1957) suggest that this behaviour belongs to the category of *attention*: this they define as 'any behaviour, motor or perceptual, which has as its end-state contact between the organism and selected portions of its environment'. The

experimental techniques which arouse attention are those which produce change: the change may be in time, as in experiments in which a rat is put in a maze, removed and then returned to it after an alteration has been made in the structure or appearance of the maze; or the change may be spatial, as in a situation in which an animal is faced with a number of separate, unfamiliar objects. But spatial change is not fundamentally different from temporal, since the differences have to be perceived by means of a scanning process which takes time. It follows also that the distinction, referred to above, between complexity and novelty, is of minor importance.

To this may be added that, whether the change is spatial or temporal, there is an 'expectation' or 'set', derived from previous experience, which is compared with the new situation; in the absence of fear, the greater the difference from expectation, the greater is the probability of approach to the new object or situation. The conclusion must be that animals will move about their living space without any impulsion from the 'homeostatic drives' such as hunger, thirst and so forth. On the one hand, novelty tends to impel investigation or, to put the point more non-committally, novelty leads to 'approach behaviour'; from this follows the notion of a 'curiosity drive': novel stimuli have 'reinforcement value' in a learning situation. On the other hand, rats (and other animals) tend to behave in such a way as to put themselves in positions in which they experience new constellations of stimuli (Berlyne, 1950). Activity persists until novelty is exhausted, or until some specific need becomes overwhelmingly strong.

IV. Avoidance behaviour

The investigatory activity of rats is so highly developed that in wild rats it is held in check by an opposite tendency, that of avoidance. Chitty & Shorten (1946), Chitty (1954) and Shorten (1954) describe observations on wild rats in the field: the essential feature of their experiments was the placing of an unfamiliar object, such as a box or a pile of food, in a place which the rats were already accustomed to visit; or a familiar object was sometimes put in a new position. They found that such objects were often completely avoided for hours or even days, and that this 'new object reaction' could greatly influence feeding behaviour. This avoidance is in contrast to the response of wild rats on being placed in an entirely unfamiliar cage, or to being suddenly given access to a new area: in both of these situations wild rats, like tame ones, explore extensively without delay, though in a leisurely manner (Barnett & Spencer, 1951; Barnett, 1958*a*). The relationship between exploration and avoidance in wild rats, its effects on feeding behaviour, and the differences between tame and wild rats have been examined in detail by Barnett (1958*b*). In his experiments rats were kept in small cages. Some wild rats avoided unfamiliar objects even though this entailed going without food for one or more days; by contrast, albino and hooded rats investigated new objects at once and continued to feed at their normal daily rate. (Although albino rats do not display 'new object reaction' or neophobia, they can of course be frightened by other types of stimuli: when they are so frightened, their exploratory behaviour is interrupted (Montgomery, 1955; Montgomery & Monkman, 1955).)

A relationship between 'curiosity' and 'fear' has often been asserted (Darwin, 1874, pp. 108 et seq.; McDougall, 1908, pp. 57–8): recently it has been discussed by Berlyne (1951) with reference to rats and man, by Hinde (1954) for the chaffinch (*Fringilla coelebs*)

and by Thorpe (1956, p. 95) in relation to habituation. Dember & Earl (1957) make no distinction between approach and avoidance for the purposes of their discussion: they point out that both involve attention.

In wild rats, as we have seen, investigation dominates in a totally new environment, while avoidance (flight, 'fear') results from a relatively minor change in a familiar constellation. The difference between a wild rat's response to total novelty on the one hand, and to a 'new object' on the other, corresponds to a distinction made by Berlyne (1951). He points out that 'change' or 'novelty' may signify either (*a*) new in relation to the individual's whole life history, or (*b*) new in relation to what has been previously experienced. It is the latter which may cause 'surprise', 'fear' and sometimes flight, instead of 'curiosity'. The important factor in the evocation of avoidance behaviour is evidently a contrast between what is expected and what is observed. In one respect the result of avoidance is the same as that of approach: the stimulus, which at first arouses an 'emotion' eventually comes no longer to do so. When the response is investigation, the stimulus gradually becomes ineffective through habituation; when it is flight, by its removal from the perceptual field. Even when flight is the first response, investigation comes later; and if the new object is food it will eventually be eaten. This response to an *unfamiliar* food is wholly different from that to a *distasteful* one: even when hunger forces a rat to eat the latter, no appetite for it develops and, as soon as opportunity offers, another food is chosen (Barnett & Spencer, 1953*a*).

Hebb (1946) provides an example of 'neophobia', similar to that of rats, in his study of fear and anger in chimpanzees. Of the stimuli which evoke these responses he writes: 'An essential feature of the stimulating conditions is the divergence of the object avoided from a familiar group of objects while still having enough of their properties to fall within the same class.' It is possible that this sort of behaviour is widespread at least among mammals. It is, however, certainly not universal. Southern (1954) has shown that in the house mouse (*Mus musculus*) it is replaced by a mercurial and unpredictable type of behaviour: and Crowcroft (1951) found no 'new object reaction' in shrews (*Sorex araneus* and *S. minutus*).

V. Exploratory behaviour in ontogeny

(1) *'Innate' or 'learned'?*

Behaviour is often described as either 'innate' ('instinctive') or 'learned', with the implication that the two categories are, on the whole, separate. Some difficulties arising from the concept of instinctive behaviour have already been mentioned (II, 1). The distinction between 'innate' and 'learned' is commonly implied to be a developmental one: while innate behaviour develops (it is supposed) in any environment which allows the animal to survive, learned behaviour depends on the specific conditions acting on the individual, and in particular acting on the animal's nervous system.

In which category, then, should spontaneous alternation, exploration and avoidance appear? In all three, the responses depend on previous experience, that is, on the experience of relationships which can be compared with the new ones; nevertheless, the response to change is made at once, without any gradual learning by practice.

It might be argued that all these instances of response to novelty are examples of insight behaviour. The latter is customarily listed as one of the categories of learned behaviour (e.g. Thorpe, 1956); it is difficult to devise a definition to cover all the various sorts of

behaviour referred to by this term, but Thorpe (1950) suggests 'the sudden production of a new adaptive response not arrived at by...trial and error behaviour'. The word 'random' precedes 'trial and error' in the original passage: it has been omitted here because it is questionable to what extent trial and error behaviour is ever random in the rigorous, statistical sense. In so far as spontaneous alternation is a general phenomenon in the animal kingdom, it must follow that truly random movements do not occur. In the definition as quoted, the word 'adaptive' seems to contain a subjective element, since it appears to suggest that the observer must approve of the response before classing it as 'insight behaviour'. It might be claimed that 'adaptive' means 'favourable to the survival of the individual or the species', but this is not an easy criterion to apply: the survival value of particular acts of behaviour is sometimes difficult to assess.

This comment on the concept and definition of 'insight behaviour' is not intended to disparage the valuable discussion and definitions contained in Thorpe's paper. The object is to discuss difficulties of a sort, already encountered in II, 1, which arise almost inevitably when one attempts formal definitions of terms used to describe complex and diverse activities. The responses to familiar and to novel stimulus situations described here have two things in common with the various sorts of 'insight behaviour': they depend on previous experience, and they are sudden. They differ, however, in taking the form always of approach ('curiosity') or avoidance (flight, 'fear'); the movements are thus always orientated on the new object or situation, and are not related to a 'goal-object' such as food or to a consummatory state such as a full belly. Exploratory behaviour does, however, achieve a 'consummatory situation' in which the whole of an area has been 'investigated' and, as will be seen further in VI, 1, the topography of it has been learnt; but this does not apply to avoidance.

It seems doubtful whether any purpose could be served by calling exploration or avoidance insight behaviour, or indeed by putting these types of response in any of the standard categories of learned behaviour. (It will be seen, however, in VI, 3, that insight learning may *accompany* exploration.) Hebb (1949, p. 245) considers that avoidance behaviour can properly be labelled neither 'innate' nor 'learned'. The same applies to exploration (and possibly to many other types of behaviour). Perhaps a new terminology is needed; but, whatever vocabulary is used, it seems that there is at present no justification for grouping all kinds of response to novelty under one head.

(2) *Learning to learn*

Whether response to novelty is called 'innate' or 'learned' or neither, it is certain that it occurs in some form very early in post-natal development. Rigorous experiments have not been done on the exploratory behaviour of infant rats but, as Small (1899) first pointed out, rats move about a good deal soon after birth; some of this movement seems to have nothing to do with seeking food or warmth, though it may be related to specific tactile stimuli. Detailed study is needed on this point. Manipulatory behaviour has been observed in wild rat nestlings: at the age of 13 days, before the eyes were open, they have been seen to take fragments of wood in both hands and, in a clumsy manner, convey them to the mouth. (Harlow *et al.* (1956) have similarly observed 'self-sustaining' manipulatory behaviour in the infant rhesus monkey.) Further, as soon as the eyes open, at about 17 days, general exploration of the environment begins, although this involves leaving the nest and the mother; within a few days the topography of an area around the

nest has been learnt, and this area is repeatedly re-explored during play (unpublished observations).

Hebb (1949, p. 298) has discussed the difference between early and late learning in mammals, and has emphasized the peculiar importance of the former. He gives evidence that rats allowed a richer experience of diverse environmental conditions early in life display greater 'intelligence' as adults: that is, they are better able to profit from new experiences than are rats brought up in a more restricted environment. The existence of this 'deutero-learning' (Bateson, 1942) in rats has now been confirmed in a number of studies. Bingham & Griffiths (1952) and Forgays & Forgays (1952) repeated Hebb's observations with a more rigorous technique, and Hymovitch (1952) showed that diversity of experience later in life has relatively little effect on learning ability; he also emphasized the importance of perceptual, as distinct from motor, learning. Gibson & Walk (1956) brought up rats from birth in cages with white walls bearing black circles and triangles These rats learned to discriminate between a circle and a triangle sooner and with fewer errors than control rats which had had an undecorated home.

Forgus (1954) found that the effects of this early perceptual learning are far from simple: giving rats in early life a complex environment, whether it could be explored or only seen, resulted in greater activity, less 'emotionality' and more variability of behaviour in an unfamiliar situation, than was shown by controls; discrimination of visual forms and spatial problem-solving were also superior in the experimental animals. In later experiments Forgus (1955) showed that early motor experience too can influence learning ability. One group of hooded rats was given much early visual and motor experience in a large and complex environment; a second group had similar visual experience but much less opportunity for motor activity. When the rats were tested, as adults, in a maze, the second group did the better of the two if the lights were kept on, while the first group did better in darkness, that is, in conditions in which learning had evidently to be based on kinaesthesia. This is a further example of the fact, already exemplified from studies of spontaneous alternation (III, 1), that learning can be related to any aspect of the input to the central nervous system, exteroceptive or proprioceptive. The work of Forgus (1956), like that of Hymovitch already quoted, also supports Hebb's hypothesis on the special importance of early deutero-learning. Forgus found that form discrimination and form generalization in learning experiments was enhanced by giving hooded rats experience of the forms used from the time the eyes opened, at 16 days, to the age of 41 days; and that similar exposure from 41 to 66 days was less effective.

The very extensive work on rats is paralleled by a smaller number of studies on other species. Thompson & Heron (1954*a*, *b*) have found similar effects of early experience on the 'intelligence' of dogs. Melzack & Thompson (1956) find that early restriction interferes with the normal development of social responses in dogs. In chimpanzees Nissen, Chow & Semmes (1951) have demonstrated the damage done to perceptual function by reducing tactual and kinaesthetic stimulation early in life. There can be no doubt that for mammals at least early deutero-learning is of prime importance.

It will be shown below (VI, 1) that exploratory behaviour is of the greatest importance in the process by which an individual adapts itself to its own particular environment. The evidence given in the preceding paragraphs suggests that, in mammals at least, diverse experience in early life, both sensory and motor, is also necessary for the full development of function in the central nervous system. It follows from this that there is a 'biological

need' for such experience: that is, early exploratory behaviour, by making possible a variety of sensory and motor experience, has survival value. If this is so, it is not surprising that such behaviour is found to be carried out 'for its own sake', and not merely in relation to satisfying one of the conventionally accepted homeostatic needs. The evidence for exploration as a 'behaviour pattern' in its own right, and for 'stimulus hunger', given in III, 1 and III, 2 above, might well have been sought as a result of inference from biological principles.

VI. Ecology of exploration and avoidance

(1) *Exploratory behaviour*

The discussion so far has been principally in terms of the stimulus conditions which evoke investigation or avoidance, with some reference to the largely unexplored field of the internal states concerned. Something further may now be said about the biological function, or survival value, of these types of behaviour. Wild rats, once they leave the nest, are subject to predation, and their principal means of avoiding predators are the use of pathways under cover, and flight to a burrow or other place of concealment. This behaviour depends on previous experience of the topography of their living space. Given this, they can run from any one point to any other, by the shortest route and in the least possible time. The behaviour of a rat violently disturbed in strange surroundings is quite different: it leaps and runs in an erratic and undirected manner, and often passes covered places of refuge to which it would resort at once if it were familiar with the conditions (unpublished observations). A rat in these circumstances does not pause to reflect. The effect of this is illustrated by the work of Elton (1953), who has brought evidence of the greater vulnerability of rats in a strange area to predation by cats.

Exploration does not, however, lead only to learning where cover and nest sites are: it is accompanied by sampling of all the edible or potable materials encountered (Barnett & Spencer, 1953*b*); and so a rat is regularly informed of accessible sources of food and water (Barnett, 1956). A further consequence of general movement about the living space is that odour trails are left on the routes regularly used, and these are detected by other rats (Reiff, 1952, 1953).

(2) *Avoidance behaviour*

If wild rats behaved as tame ones do towards unfamiliar objects, their compulsive exploratory and sampling behaviour would quickly lead them into traps or result in the ingestion of poison bait. 'New object reaction' or neophobia, displayed in a highly developed form by wild rats, protects them from the consequences of 'curiosity'; and the studies, already quoted, of Chitty & Shorten (1946), Chitty (1954) and Shorten (1954) have shown how effective this behaviour is in preserving rats from human efforts to kill them. It seems likely that the extent to which wild rats commensal with man display avoidance behaviour is a consequence of selection over the 7000 years during which civilization, and therefore stores of human food, have existed. (More intensive selection *against* wildness and neophobia in laboratory rats has been completely successful in less than a century.)

The neophobia of wild rats is not by itself sufficient protection against poisoned food. It is combined with a capacity to learn to refuse toxic mixtures (Rzoska, 1953, 1954). This parallels the well-known ability to select, in certain instances, the nutritionally superior of two foods: such dietary self-selection (reviewed by Munn, 1950) is in part a result of

associative learning. We have seen that a wild rat regularly samples everything edible within range, although unfamiliar foods are at first avoided and then provisionally sampled only in very small amounts. This combination of exploring and avoidance with learning is elegantly adapted to giving a rat the maximum of information about the resources and dangers of its environment, in the safest possible way (Barnett, 1956).

(3) *Exploratory (latent) learning*

Wilkinson (1952) has given a mathematical analysis of the effects of strictly random exploration; his concern was with the 'navigation' of birds, but the analysis has a quite general application, and demonstrates that random searching is a remarkably effective method of reaching a specific goal. But the studies of spontaneous alternation show that exploratory behaviour is not random: on the contrary, areas which have just been explored tend to be avoided; this must make exploration correspondingly more efficient, either in leading the animal to a particular goal such as food, or in making it generally familiar with a large area. Further, most of the movements of animals are directional, for instance, to or away from light (phototaxis). Rats are specifically influenced by tactile stimuli, as well as light, and tend to move in contact with vertical surfaces. This further reduces the randomness of their movements. It also tends to keep them under cover. This tendency is, however, not strong enough to prevent them from exploring, eventually, all parts of their home range.

Whether random or not, exploration would be largely ineffectual without a concomitant learning process. Accordingly, in the preceding sections it has been taken for granted that exploratory behaviour is accompanied by topographical learning. Certainly, it is obvious that wild animals rapidly become acquainted with the topography of a newly colonized area. But the way in which such learning is related to exploration has been rigorously established in laboratory studies of tame rats.

Just as exploratory behaviour is performed 'without motivation', so apparently unmotivated learning takes place while it is being carried out. This *exploratory* (latent) *learning* has been defined as 'the association of indifferent stimuli or situations without patent reward' (quoted by Thorpe, 1956, p. 91). This subject has been so fully reviewed by Munn (1950, pp. 315 et seq. and 401 et seq.), by Thistlethwaite (1951) and Thorpe (1956), that only a brief account is needed here.

The essential features of this behaviour were first made clear by Blodgett (1929). If rats are given experience of a maze, but are not rewarded on reaching the goal point, they explore the maze. If, later, they are run in the same maze in the usual way, with a reward of food at the goal point, they learn the maze more quickly than controls which have had no unrewarded experience of the maze: they make fewer errors and their running time is shorter.

This apparently 'unrewarded' learning was received with scepticism by some psychologists who were wedded to the 'law of reinforcement', according to which learning takes place only when the animal is 'motivated'. Hull (1943), in attempting to make the concept of motivation more concrete, equated it with a state of 'primary need', but Hebb (1949, pp. 176 et seq.) and Hebb & Mahut (1955) have effectively criticized this formulation. As far as the specific case of exploratory learning is concerned, many workers have confirmed Blodgett's observations (see the reviews already quoted). In some experiments, however, exploratory learning did not take place when it might have been expected.

These were studies in which the rats were not satiated during their preliminary experiences of the maze, but were in a state of 'irrelevant drive'. For example, Spence & Lippitt (1946) trained rats to run a Y-maze while thirsty; one path led always to water, the other to food for half the rats and to nothing for the other half; all the rats were able to explore the maze. When learning was established, the rats were run while hungry but not thirsty; and food was put in the, formerly empty, goal box of the control group. In these conditions, no difference was found between the two groups in speed of learning to go to food, although it might be expected that those rats which had had experience of food in the goal box when not hungry would have learned the task sooner. The evidence from such experiments suggests that generalized exploratory learning takes place only in the *absence* of special need, for example for food or water: when an animal is severely deprived, it learns only those responses which lead to the relief of its particular need.

It was shown in II, 3 that the variable movements of rats (appetitive behaviour in the general sense) may be related to a specific need or may be performed 'for their own sake'. Similarly, as the preceding paragraphs show, the learning that accompanies movement about the living space is of two kinds. First, there is the learning which results from encountering a specific goal object (food, water, nest site, nest materials or other rats); the same type of learning occurs if a noxious stimulus, such as harmful food, is experienced. Secondly, the animal is all the time liable to be influenced by the non-specific stimuli of its environment. Exploratory behaviour occurs at all ages in both tame and wild rats (Barnett, 1956), and it must be supposed that the central nervous system, especially the cerebral cortex, is being continually modified to some degree by the afferent impulses reaching it; this effect is evidently most generalized when the animal is not in a state of primary need. It is comparable to, but more specific than, the deutero-learning referred to in V, 2.

VII. Exploration and theories of behaviour

(1) *Learning*

The study of exploratory behaviour has influenced the ways in which certain aspects of behaviour are interpreted. First, it has been used to throw light on the nature of complex learned behaviour.

The topographical learning that accompanies variable movements about the living space, whether related to some specific state of need or not, is exceedingly complex. This topic (reviewed by Thorpe, 1956, pp. 91 et seq., 109 et seq.) is still controversial. One of the problems is the extent to which this behaviour can be regarded as the learning of a sequence of motor acts which lead the animal through a succession of positions from the starting point to the goal. An interpretation in these terms is sometimes called a response theory, or *SR* theory, of learning. It stems from the notion, long discarded by most students of animal behaviour, that behaviour can be adequately described as if it consisted entirely of 'reflexes' and 'conditional reflexes'. It has been known for more than thirty years that no such theory can account for all the behaviour shown in mazes or other situations requiring problem-solving, because animals can use different series of movements, on different occasions after a task has been learnt, to solve a problem. The early work on this subject was reviewed with admirable lucidity by Lashley (1929, 1934). This behaviour comes within the category of insight behaviour as defined in V, 1. Many experiments have been carried out, since Lashley wrote, confirming that rats can profit

from separate previous experiences in taking the shortest route to a goal, and that in doing so they may follow a path which they have never employed before. Tolman (1948), in a review of this work, writes that it is as if 'in the course of learning something like a field map of the environment gets established in the rat's brain' (see also Caldwell & Jones, 1954). Tolman uses the term 'expectancy' to describe the state of a rat when it is behaving in this way.

Very recently, Birch & Korn (1958) have done further experiments to test Tolman's idea of 'cognitive maps', and have subjected the various theoretical positions in this area to critical scrutiny. They point out that, even though simple stimulus-response theories of topographical learning are inadequate, it does not follow from this that the behaviour can be explained only by a 'cognitive' or 'expectancy' theory. Their experiments, carried out in apparatus similar to that used by Tolman and his colleagues, suggest that, when a rat seems to be behaving as if it possessed a cognitive map of a maze, it is in fact responding to particular stimuli, such as differences of illumination, which are associated with food; they further find that part of the series of responses depends on kinaesthetic cues. Their position then includes a 'response' element, in that proprioceptive stimuli are held to guide the rat in some of its movements. But other movements may be directed towards the goal, as a result of previous learning of the position of the goal in relation to sources of stimulation; and these movements are not mere repetitions of those previously learned in a different situation.

The discussion is unlikely to stand still at this point. So far, it has effectively shown that very simple theories of topographical learning are untenable, and it has at least displayed some of the difficulties of constructing a satisfactory one.

(2) *Motivation*

The second theoretical aspect of the study of exploratory behaviour concerns the internal states which make an animal active: that is, it concerns motivation. Here it is desirable to refer again to the two distinct ways in which behaviour may be 'explained'. First, an activity may be said to have survival value for the individual or species in which it occurs, and so be explained by the action of natural selection; the causes of the behaviour, implied in any such statement, are then events in the comparatively remote past which have acted in the first place on ancestors of the animals observed. Secondly, an activity may be explained as being due to stimuli influencing the animal at the time, or at least during its own life history; or as due to internal processes, for instance in the nervous system.

The first type of explanation is exemplified in VI, 1 to VI, 3: exploratory and avoidance behaviour, and the accompanying learning, have obvious survival value. Many of the easily observed types of behaviour confer some evident advantage on the individual or on the species. This has led to the notion of the homeostatic function of behaviour: an animal's activities are looked upon as all tending to fulfill its primary needs (including those connected with reproduction). Hence, the internal state or motivation responsible for specific activity has been equated with a state of specific need, notably by Hull (1943). However, the fact that behaviour is a product of natural selection, and tends to ensure survival, does not necessarily imply that particular activities are always a direct consequence of the needs they help to satisfy. To say that they are, is to confuse the two types of cause defined above. A kind of behaviour may tend to satisfy a need, and yet be per-

formed regardless of the presence or absence of the need: it may be evoked by internal processes which are independent of such states as hunger or thirst, and by external stimuli which are neutral as far as survival is concerned.

This is the case with exploratory behaviour. Exploration therefore obviously demands a theory of motivation more subtle and complex than those which have prevailed hitherto. It has been shown above (VII, 1) that stimulus-response theories are inadequate to account for exploratory learning. Such theories, based on a misleadingly simple account of conditional reflexes, have also been unable to cope with the obvious facts of motivation, such as the 'spontaneous' fluctuations in an animal's response to a given stimulus or situation. Accordingly, emphasis has come to be laid on autonomous changes in the central nervous system (e.g. Young, 1938; Hebb, 1949).

It is now evident that exploratory behaviour and learning depend on yet a further feature of the total situation, namely, non-specific changes in the stimulus pattern falling on the sense organs. By 'non-specific changes' is meant those which do not necessarily evoke or demand any special consummatory response. General exteroceptive stimulation has long been thought to be important in maintaining central nervous activity in, for example, insects (Wolsky, 1933). The work on the sense organs of invertebrates, with which Wolsky was chiefly concerned, has however not demonstrated any special significance in a *varying* level of stimulation. Nevertheless, Darchen (1952, 1955) has shown that cockroaches (*Blattella germanica*) display exploratory behaviour like that of rats; and it may be that *Stimulationsorgane* are of importance as channels for external stimulation which is needed to maintain this behaviour. This requires investigation.

In mammals such as rats or monkeys the change in stimulation which evokes exploration may result from the presence of objects which are themselves changing; or it may be due to the scanning, by the animal, of an environment containing a variety of objects. This variety both stimulates activity and satisfies a need. In this second quality, though not in other respects, the condition of being stimulated resembles a consummatory state. In II, 1 above, 'consummatory state' was by implication defined as an afferent pattern, impinging on the central nervous system, which reduces or ends an activity. But an important feature of a consummatory state is that it has reward value: an animal tends to learn to achieve it. This is the case too with variety of stimulation (III, 2): the opportunity to experience sensory change alone can act as an incentive to learn some task; but, unlike a consummatory state, it leads to greater activity until the novelty wears off.

Hebb (1955) accordingly distinguishes two types of effect resulting from a sensory event. First, there is the *cue function*, which is the familiar one of guiding behaviour, that is, of determining that a particular behaviour pattern will be performed and orientated in an appropriate direction. Secondly, there is the *arousal function*. The importance of this second function is attested by the evidence already given in the present review. Hebb suggests that the state of arousal, which depends on varying stimulation, is synonymous with a 'general drive state', and that this state is maintained in mammals by the reticular formation of the brain stem: on this view, 'the conception of drive assumes anatomical and physiological identity'. The internal springs of activity which we call motivation are consequently dependent to an important extent on the changing stimulation of the exteroceptors.

VIII. Summary

Many animals have a general tendency to explore their surroundings, that is, to approach and to enter every accessible place within a considerable area around a nest or resting place. This behaviour has been especially well studied in rats (*Rattus norvegicus*), both tame and wild. Exploratory behaviour is one category of the *variable* movements of animals. Often, such movements end when a specific (consummatory) state has been achieved (whether by performance of a consummatory act or not). Sometimes the variable activity ends without attainment of any recognizable consummatory state. In either case, it is often possible to assert, or to infer, that the activity has been *evoked* by a specific internal state (such as hunger). All these diverse activities may be labelled 'appetitive'; this term then distinguishes variable movements from stereotyped action patterns such as courtship ceremonies; the latter are, by comparison, not variable but fixed (predictable in detail) within each species.

The definition of 'appetitive behaviour' thus becomes 'any variable behaviour'. Consequently, it is necessary to state for each instance to what 'goal object' (if any) and to what internal state (if known) it is related. Some exploratory behaviour is, however, unrelated to any goal, and independent of any special internal state apart from that which brings about the exploration itself.

Exploratory behaviour in satiated rats provides a clear example of variable movements unrelated to specific need. It is observed most readily in the laboratory varieties, but is highly developed also in wild rats. It is not a spontaneous motor activity, but a result of a general tendency to behave so as to vary the stimulus situation experienced by the animal from moment to moment, and to avoid what has just been experienced. Other species, from cockroaches to monkeys, behave in a similar way.

'Spontaneous alternation' is a special case of exploratory behaviour. It is not, as has been suggested, a result of inhibition of the motor *response* which has just been performed: in most experimental situations what is avoided is repetition of visual experiences, but it is possible to devise experiments in which response alternation occurs. All experimental data on spontaneous alternation may be subsumed under the heading of '*stimulus satiation*'. Similarly, all exploratory behaviour may be interpreted in terms of 'stimulus hunger'. More generally, it may be categorized as *attention to change*. However, *activity* may be increased, not only by changes in the stimulus situation, but also by deprivation, for instance of food: this (*a*) lowers the threshold of response to external change, (*b*) may directly increase activity.

Wild rats given access to an unfamiliar area explore it. They also regularly re-explore their home range. But movement into a familiar area is inhibited by the presence of an unfamiliar object or arrangement. This new object reaction, or neophobia, is not found in tame rats. Its biological value is to protect rats from danger. Similar responses occur in some other species.

Exploratory behaviour begins in early infancy but, having no specific pattern, it differs greatly from behaviour of the kinds commonly called 'innate'. It depends on learning, since it is a consequence of comparing a stimulus pattern with one previously experienced; but it does not fall into any of the conventional classes of learned behaviour. Like avoidance behaviour, it exemplifies the inadequacy of the division of behaviour into 'innate' and 'learned'.

Visual and motor exploration of a complex habitat during early life is necessary for the full development of a mammal's ability to solve problems. Early experience is accompanied by a process of 'deutero-learning', or learning to learn, which occurs little or not at all in the adult.

The survival value of exploration lies not only in the consequent deutero-learning, but also in the learning and relearning of the resources of the animal's environment which accompanies it. In rats this learning, combined with the avoidance of new objects, gives the maximum of information about the environment, in the safest possible way. This exploratory, or latent, learning is 'unrewarded', and its effects become evident only when the animal is later faced with a specific problem. It is a highly complex form of learning, and gives rise to a form of insight behaviour, since it enables an animal to make use of two or more previous experiences to perform a new action.

The study of exploratory behaviour has thrown new light on 'motivation', since it has shown that exteroceptive stimulation has a necessary arousal function which influences the animal's general level of activity.

References

Adlerstein, A. & Fehrer, E. (1955). The effect of food deprivation on exploratory behaviour in a complex maze. *J. Comp. Physiol. Psychol.* **48**, 250–3.

Barnett, S. A. (1955). 'Displacement' behaviour and 'psychosomatic' disorder. *Lancet*, **269**, 1203–8.

Barnett, S. A. (1956). Behaviour components in the feeding of wild and laboratory rats. *Behaviour*, **9**, 24–43.

Barnett, S. A. (1958*a*). An analysis of social behaviour in wild rats. *Proc. Zool. Soc. Lond.* **130**, 107–52.

Barnett, S. A. (1958*b*). Experiments on 'neophobia' in wild and laboratory rats. *Brit. J. Psychol.* **49**, 195–201.

Barnett, S. A. & Spencer, M. M. (1951). Feeding, social behaviour and inter-specific competition in wild rats. *Behaviour*, **3**, 229–42.

Barnett, S. A. & Spencer, M. M. (1953*a*). Responses of wild rats to offensive smells and tastes. *Brit. J. Anim. Behav.* **1**, 32–7.

Barnett, S. A. & Spencer, M. M. (1953*b*). Experiments on the food preference of wild rats. *J. Hyg. Camb.*, **51**, 16–34.

Bastock, M., Morris, D. & Moynihan, M. (1953). Some comments on conflict and thwarting in animals. *Behaviour*, **6**, 66–84.

Bateson, G. (1942). Social planning and the concept of deutero-learning. In *Conference on Science, Philosophy and Religion, Second Symposium*. New York.

Beach, F. A. (1951). Instinctive behavior: reproductive activities. In *Handbook of Experimental Psychology*. New York.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.* **41**, 68–80.

Berlyne, D. E. (1951). Attention to change. *Brit. J. Psychol.* **42**, 269–78.

Berlyne, D. E. (1955). The arousal and satiation of perceptual curiosity in the rat. *J. Comp. Physiol. Psychol.* **48**, 238–46.

Berlyne, D. E. & Slater, J. (1957). Perceptual curiosity, exploratory behavior, and maze learning. *J. Comp. Physiol. Psychol.* **50**, 228–32.

Bexton, W. H., Heron, W. & Scott, T. H. (1954). Effects of decreased variation in the sensory environment. *Canad. J. Psychol.* **8**, 70–6.

Bingham, W. E. & Griffiths, W. J. (1952). The effect of different environments during infancy on adult behavior in the rat. *J. Comp. Physiol. Psychol.* **45**, 307–12.

Birch, H. G. & Korn, S. J. (1958). Place-learning, cognitive maps, and parsimony. *J. Gen. Psychol.* **58**, 17–35.

Blodgett, H. C. (1929). The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.* **4**, 113–34.

Butler, R. A. (1953). Discrimination learning by rhesus monkeys to visual-exploration motivation. *J. Comp. Physiol. Psychol.* **46**, 95–8.

Butler, R. A. (1954). Incentive conditions which influence visual exploration. *J. Exp. Psychol.* **48**, 19–23.

Butler, R. A. (1957). Discrimination learning by rhesus monkeys to auditory incentives. *J. Comp. Physiol. Psychol.* **50**, 239–41.

Butler, R. A. & Alexander, H. M. (1955). Daily patterns of visual exploratory behavior in the monkey. *J. Comp. Physiol. Psychol.* **48**, 247–9.

Butler, R. A. & Harlow, H. F. (1954). Persistence of visual exploration in monkeys. *J. Comp. Physiol. Psychol.* **47**, 258–63.

Buytendijk, F. J. J. (1931). Eine Methode zur Beobachtung von Ratten in aufgabefreien Situationen. *Arch. néerl. Physiol.* **16**, 574–96.

Caldwell, W. E. & Jones, H. B. (1954). Some positive results on a modified Tolman and Honzik insight maze. *J. Comp. Physiol. Psychol.* **47**, 416–8.

Calhoun, J. B. (1949). A method for self control of population growth among mammals living in the wild. *Science*, **109**, 333–5.

Campbell, B. A. & Sheffield, F. D. (1953). Relation of random activity to food deprivation. *J. Comp. Physiol. Psychol.* **46**, 320–2.

Chapman, R. M. & Levy, N. (1957). Hunger drive and reinforcing effect of novel stimuli. *J. Comp. Physiol. Psychol.* **50**, 233–8.

Chitty, D. (1954). The study of the brown rat and its control by poison. In *The Control of Rats and Mice*, vol. I. Oxford.

Chitty, D. & Shorten, M. (1946). Techniques for the study of the Norway rat (*Rattus norvegicus*). *J. Mammal.* **27**, 63–78.

Crowcroft, P. (1951). Keeping British shrews in captivity. *J. Mammal.* **32**, 354–5.

Darchen, R. (1952). Sur l'activité exploratrice de *Blattella germanica*. *Z. Tierpsychol.* **9**, 362–72.

Darchen, R. (1955). Stimuli nouveaux et tendance exploratrice chez *Blattella germanica*. *Z. Tierpsychol.* **12**, 1–11.

Darwin, C. (1874). *The Descent of Man*, 2nd edition. London.

Dember, W. N. (1956). Response by the rat to environmental change. *J. Comp. Physiol. Psychol.* **49**, 93–5.

Dember, W. N. & Earl, R. W. (1957). Analysis of exploratory, manipulatory and curiosity behaviors. *Psychol. Rev.* **64**, 91–6.

Dember, W. N., Earl, R. W. & Paradise, N. (1957). Response by rats to differential stimulus complexity. *J. Comp. Physiol. Psychol.* **50**, 514–8.

Eayrs, J. T. (1954). Spontaneous activity in the rat. *Brit. J. Anim. Behav.* **2**, 25–30.

Elton, C. (1953). The use of cats in farm rat control. *Brit. J. Anim. Behav.* **1**, 151–5.

Fehrer, E. (1956). The effects of hunger and familiarity of locale on exploration. *J. Comp. Physiol. Psychol.* **49**, 549–52.

Finger, F. W. & Reid, L. S. (1952). The effect of water deprivation and subsequent satiation upon general activity in the rat. *J. Comp. Physiol. Psychol.* **45**, 368–72.

Forgays, D. G. & Forgays, J. W. (1952). The nature of the effect of free-environmental experience in the rat. *J. Comp. Physiol. Psychol.* **45**, 322–8.

Forgus, R. H. (1954). The effect of early perceptual learning on the behavioral organization of adult rats. *J. Comp. Physiol. Psychol.* **47**, 331–6.

Forgus, R. H. (1955). Early visual and motor experience as determiners of complex maze-learning ability under rich and reduced stimulation. *J. Comp. Physiol. Psychol.* **48**, 215–20.

Forgus, R. H. (1956). Advantage of early over late perceptual experience in improving form discrimination. *Canad. J. Psychol.* **10**, 147–55.

Gibson, E. J. & Walk, R. D. (1956). The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *J. Comp. Physiol. Psychol.* **49**, 239–42.

Glanzer, M. (1953*a*). The role of stimulus satiation in spontaneous alternation. *J. Exp. Psychol.* **25**, 387–43.

Glanzer, M. (1953*b*). Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.* **60**, 257–68.

Hall, J. F. (1956). The relationship between external stimulation, food deprivation, and activity. *J. Comp. Physiol. Psychol.* **49**, 339–41.

Hardy, J. D. (1954). Control of heat loss and heat production in physiologic temperature regulation. *Harvey Lect.* **49**, 242–70.

Harlow, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *J. Comp. Physiol. Psychol* **43**, 289–94.

Harlow, H. F., Blazek, N. C. & McClearn, G. E. (1956). Manipulatory motivation in the infant rhesus monkey. *J. Comp. Physiol. Psychol.* **49**, 444–8.

Harlow, H. F., Harlow, M. K. & Meyer, D. R. (1950). Learning motivated by a manipulation drive. *J. Exp. Psychol.* **40**, 228–34.

Harlow, H. F. & McClearn, G. E. (1954). Object discrimination learned by monkeys on the basis of manipulation motives. *J. Comp. Physiol. Psychol.* **47**, 73–6.

Havelka, J. (1956). Problem-seeking behaviour in rats. *Canad. J. Psychol.* **10**, 91–7.

Heathers, G. L. (1940). The avoidance of repetition of a maze reaction in the rat as a function of the time interval between trials. *J. Psychol.* **10**, 359–80.

Hebb, D. O. (1946). On the nature of fear. *Psychol. Rev.* **53**, 259–76.

Hebb, D. O. (1949). *The Organization of Behavior*. New York.

Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.* **62**, 243–54.

Hebb, D. O. & Mahut, H. (1955). Motivation et recherche du changement perceptif chez le rat et chez l'homme. *J. Psychol. norm. path.* **52**, 209–21.

Hediger, H. (1950). *Wild Animals in Captivity*. London.

Hinde, R. A. (1953). Appetitive behaviour, consummatory act and the hierarchical organization of behaviour—with special reference to the Great Tit (*Parus major*). *Behaviour*, **5**, 189–224.

HINDE, R. A. (1954). Factors governing the changes in strength of a partially inborn response. 1. The nature of the response, and an examination of its course. *Proc. Roy. Soc.* B, **142**, 306–31.

HULL, C. L. (1943). *Principles of Behavior.* New York.

HULL, C. L. (1951). *Essentials of Behavior.* New Haven.

HYMOVITCH, B. (1952). The effects of experimental variations on problem solving in the rat. *J. Comp. Physiol. Psychol.* **45**, 313–21.

IERSEL, J. J. A. VAN (1953). An analysis of the parental behaviour of the male three-spined stickelback (*Gasterosteus aculeatus* L.). *Behaviour Suppl.* No. 3.

JEROME, E. A., MOODY, J. A., CONNOR, T. J. & FERNANDEZ, M. B. (1957). Learning in a multiple-door situation under various drive states. *J. Comp. Physiol. Psychol.* **50**, 588–91.

KORTLANDT, A. (1955). Aspects and prospects of the concept of instinct. *Arch néerl. Zool.* **11**, 154–284.

KRECHEVSKY, I. (1937). Brain mechanisms and variability. II. Variability where no learning is involved. *J. Comp. Psychol.* **23**, 139–59.

LASHLEY, K. S. (1929). *Brain Mechanisms and Intelligence.* Chicago.

LASHLEY, K. S. (1934). Nervous mechanisms in learning. In *Handbook of General Experimental Psychology.* Worcester, Massachusetts.

LEHRMAN, D. S. (1953). A critique of Konrad Lorenz's theory of instinctive behavior. *Quart. Rev. Biol.* **28**, 337–63.

LINDAUER, M. (1952). Ein Beitrag zur Frage der Arbeitsteilung im Bienenstaat. *Z. vergl. Physiol.* **34**, 299–345.

LOCKE, N. M. (1936). A preliminary study of a social drive in the white rat. *J. Psychol.* **1**, 255–60.

LORENZ, K. (1954). Plays and vacuum activities. In *L'instinct dans le comportement des animaux et de l'homme.* Paris.

MCDOUGALL, W. (1908). *An Introduction to Social Psychology.* London.

MAJORANA, A. (1950). Ricerche sull'apprendimento dei ratti in labirinto. *Riv. Psicol.* **46**, 233–51.

MELZACK, R. & THOMPSON, W. R. (1956). Effects of early experience on social behavior. *Canad. J. Psychol.* **10**, 82–90.

MONTGOMERY, K. C. (1951). The relation between exploratory behavior and spontaneous alternation in the white rat. *J. Comp. Physiol. Psychol.* **44**, 582–9.

MONTGOMERY, K. C. (1952*a*). Exploratory behavior and its relation to spontaneous alternation in a series of maze exposures. *J. Comp. Physiol. Psychol.* **45**, 50–7.

MONTGOMERY, K. C. (1952*b*). A test of two explanations of spontaneous alternation. *J. Comp. Physiol. Psychol.* **45**, 287–93.

MONTGOMERY, K. C. (1953*a*). Exploratory behavior as a function of 'similarity' of stimulus situations. *J. Comp. Physiol. Psychol.* **46**, 129–33.

MONTGOMERY, K. C. (1953*b*). The effect of the hunger and thirst drives upon exploratory behavior. *J. Comp. Physiol. Psychol.* **46**, 315–9.

MONTGOMERY, K. C. (1953*c*). The effect of activity deprivation upon exploratory behavior. *J. Comp. Physiol. Psychol.* **46**, 438–41.

MONTGOMERY, K. C. (1954). The role of the exploratory drive in learning. *J. Comp. Physiol. Psychol.* **47**, 60–4.

MONTGOMERY, K. C. (1955). The relation between fear induced by novel stimulation and exploratory behavior. *J. Comp. Physiol. Psychol.* **48**, 254–60.

MONTGOMERY, K. C. & MONKMAN, J. A. (1955). The relation between fear and exploratory behavior. *J. Comp. Physiol. Psychol.* **48**, 132–6.

MOYNIHAN, N. (1953). Some displacement activities of the blackheaded gull. *Behaviour*, **5**, 58–80.

MUNN, N. L. (1950). *Handbook of Psychological Research on the Rat.* New York.

MYERS, A. K. & MILLER, N. E. (1954). Failure to find a learned drive based on hunger; evidence for learning motivated by 'exploration'. *J. Comp. Physiol. Psychol.* **47**, 528–36.

NISSEN, H. W., CHOW, K. L. & SEMMES, J. (1951). Effects of restricted opportunity for tactual, kinesthetic, and manipulative experience on the behavior of the chimpanzee. *Amer. J. Psychol.* **64**, 485–507.

REIFF, M. (1952). Über Territoriumsmarkierung bei Hausratten und Hausmäusen. *Verh. schweiz. naturf. Ges. Luzern*, 1951, 150–1.

REIFF, M. (1953). Differenzierung im ökologischen Verhalten bei Wanderrattenpopulationen. *Rev. suisse Zool.* **60**, 447–52.

RICHTER, C. P. (1927). Animal behavior and internal drives. *Quart. Rev. Biol.* **2**, 307–43.

RZOSKA, J. (1953). Bait shyness, a study in rat behaviour. *Brit. J. Anim. Behav.* **1**, 128–35.

Rzoska, J. (1954). The behaviour of white rats towards poison baits. In *The Control of Rats and Mice*, vol. II. Oxford.

Seitz, A. (1940). Die Paarbildung bei einigen Cichliden. 2. Die Paarbildung bei *Astatotilapia strigigena*. *Z. Tierpsychol.* **5**, 74–101.

Shorten, M. (1954). The reaction of the brown rat towards changes in its environment. In *The Control of Rats and Mice*, vol. II. Oxford.

Small, W. S. (1899). Notes on the psychic development of the young white rat. *Amer. J. Psychol.* **11**, 80–100.

Solomon, R. L. (1948). The influence of work on behavior. *Psychol. Bull.* **45**, 1–40.

Southern, H. N. (1954). *The Control of Rats and Mice*, vol. III. Oxford.

Spence, K. W. & Lippitt, R. (1946). An experimental test of the sign-gestalt theory of trial and error learning. *J. Exp. Psychol.* **36**, 491–502.

Strong, P. N. (1957). Activity in the white rat as a function of apparatus and hunger. *J. Comp. Physiol. Psychol.* **50**, 596–600.

Sutherland, N. S. (1957). Spontaneous alternation and stimulus avoidance. *J. Comp. Physiol. Psychol.* **50**, 358–62.

Thistlethwaite, D. (1951). A critical review of latent learning and related experiments. *Psychol. Bull.* **48**, 97–129.

Thompson, W. R. & Heron, W. (1954*a*). The effect of restricted early experience on the problem-solving capacity of dogs. *Canad. J. Psychol.* **8**, 17–31.

Thompson, W. R. & Heron, W. (1954*b*). The effects of early restriction on activity in dogs. *J. Comp. Physiol. Psychol.* **47**, 77–82.

Thompson, W. R. & Solomon, L. M. (1954). Spontaneous pattern discrimination in the rat. *J. Comp. Physiol. Psychol.* **47**, 104–7.

Thorpe, W. H. (1944). Some problems of animal learning. *Proc. Linn. Soc. Lond.* **156**, 70–83.

Thorpe, W. H. (1950). The concepts of learning and their relation to those of instinct. *Symp. Soc. Exp. Biol.* **4**, 387–408.

Thorpe, W. H. (1951). The definition of terms used in animal behaviour studies. *Bull. Anim. Behav.* **3**, 34–40.

Thorpe, W. H. (1954*a*). Some concepts of ethology. *Nature, Lond.* **174**, 101–5.

Thorpe, W. H. (1954*b*). The process of song learning in the chaffinch as studied by mean of the sound spectrograph. *Nature, Lond.*, **173**, 465–9.

Thorpe, W. H. (1956). *Learning and Instinct in Animals*. London.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208.

Walker, E. L., Dember, W. N., Earl, R. W., Fawl, C. L. & Karoly, A. J. (1955). Choice alternation: 3. Response intensity vs. response discriminability. *J. Comp. Physiol. Psychol.* **48**, 80–5.

Walker, E. L., Dember, W. N., Earl R. W. & Karoly, A. J. (1955). Choice alternation: 1. Stimulus vs. place vs. response. *J. Comp. Physiol. Psychol.* **48**, 19–23.

Welker, W. I. (1956*a*). Some determinants of play and exploration in chimpanzees. *J. Comp. Physiol. Psychol.* **49**, 84–9.

Welker, W. I. (1956*b*). Variability of play and exploratory behavior in chimpanzees. *J. Comp. Physiol. Psychol.* **49**, 181–5.

Wilkinson, D. H. (1952). The random element in bird 'navigation'. *J. Exp. Biol.* **29**, 532–60.

Wolsky, A. (1933). Stimulationsorgane. *Biol. Rev.* **8**, 370–417.

Young, J. Z. (1938). The evolution of the nervous system and of the relationship of organism and environment. In *Evolution*. Oxford.

Zimbardo, P. G. & Montgomery, K. C. (1957). The relative strengths of consummatory responses in hunger, thirst, and exploratory drive. *J. Comp. Physiol. Psychol.* **50**, 504–8.

(*Manuscript received* 16 *April* 1957)

THE ATTENTION-GETTING VALUE OF STRUCTURAL CHANGE

By ROBERT T. GREEN
University College London

The structure of a list of items in a serial learning task seems to have a reliable influence upon the probability of recalling a particular item. The results pose serious difficulties for either the gestalt or the usual behaviouristic serial interference interpretations. In particular, it would appear that earlier experiments which were explained in terms of Von Restorff's 'isolation' effect have overlooked and confounded the variable of structural change.

I. Introduction

Earlier experiments by Saul & Osgood (1950) and Green (1956, 1958) demonstrated that isolation in a temporal series has no effect on recall probabilities. Other published data seeming to support the isolation hypothesis are demonstrably contaminated by factors such as serial interference—a factor used by Gibson (1940) to generate Von Restorff's (1933) results. More recent work, not using paired associates, demands an explanation in terms of either pro-active interference or the attention-getting properties of structural change within the series. Until now it has not been possible to choose between these two hypotheses because an item of a new type has always been used to introduce the structural change, and has therefore not been subject to pro-active interference. In the present study an attempt is made to find a crucial experiment to discriminate between these two hypotheses and to demonstrate that the effectiveness of structural change as a factor favouring recall is due to its attention-getting properties.

II. Procedure

In order to resolve the confusion between pro-active interference and structural change it was decided to make *repetition* of a critical item the structural change. Although the subjects were instructed to learn a letter sequence, they were, in fact, tested on recall of the colour of a critical item. This incidental learning design was necessitated by the fact that repetition of the critical item must favour its recall as a letter. Two groups were used, and reference to the structure of the two corresponding lists in Table 1 will help to make the experimental design clearer. The lists are identical except that for the beta group the blue

Table 1. *Structure of lists*

Serial position	List alpha		List beta		
1	*Q*	Red	*Q*		
2	*K*	Green	*K*		
3	*S*	Yellow	*S*		
4	*G*	Black	*G*		
5	*A*	Green	*A*		
6	*B*	Blue	*H*		
7	*J*	Red	*J*		
8	*N*	Black	*N*		
9	*H*	Yellow	*H*	Critical item	
10	*U*	Brown	*U*		
11	*M*	Orange	*M*		
12	*H*	Blue	*B*		
13	*T*	Mauve	*T*		
14	*D*	Orange	*D*		
15	*L*	Brown	*L*		
16	*P*	Mauve	*P*		

letter *H* is exchanged in serial position with letter *B* of the same colour. In both lists the critical item is the yellow letter *H* in position 9. For the alpha group this is the first letter *H*, while for the beta group it is the second. According to our hypothesis the second letter *H* has attention-getting value because repetition in these series is a structural change. So it is predicted that members of the beta group will pay more attention to the critical item than those of the alpha group. Because of the greater attention accorded to this item they are more likely to learn its colour quite incidentally. Hence it follows that the probability of recalling the colour of the critical item should be greater under beta conditions.

Instructions to the subjects followed closely those reported in the earlier paper. No mention was made of the colour variable, the subjects being 'set' to learn letters. Each item was presented for 3 sec., followed immediately by the next, by means of a film-strip projector behind a translucent 18 inch square screen in a semi-darkened room. Immediately after the single learning trial the following instructions were read to the subjects.

'The letter *H* occurred twice in that series. Write down the colours of both the first and second letters *H* in the correct order. That is, write down the colour of the first letter *H*, labelling it 'first', and then directly below write the colour of the second letter *H*, labelling it 'second'. If you are unable to remember one or the other, simply put a cross in the appropriate position. If you can recall neither, put two crosses.' (20 sec. pause for writing.) 'When you have done this, indicate how sure you are of either judgement in the following manner. If you are certain to the extent that you would be prepared to swear to it in a court of law put a letter *C* for certain against the colour you are sure of. If you are fairly sure, but not perfectly sure put the letters *FS* against the appropriate colour. If you are quite doubtful about one or the other colour put the letter *D* beside it.'

The subjects were young adults, nearly all males, attending a Technical College. The use of a three-point scale was intended to refine the results, it being thought that not only should a greater proportion of the beta group recall the colour correctly, but that they should also be more sure of their choice. By asking for recall of both the letters *H* it was possible to see to what extent differences might be due to the subjects allocating the correct colours to the wrong position. The colours of yellow and blue were chosen to take advantage of the low incidence of tritanopia among the general population. No tests for colour defects were made, a haphazard allocation of subjects to the two groups being left to distribute what small effects might exist.

III. Results and Discussion

The main findings are set out in Table 2. For the alpha group the critical data are the yellow responses for the first letter *H*, and for the beta group the yellow responses for the second letter *H*. These totals, ignoring the rating scale, yield recall probabilities of 0·3

Table 2. *Analysis of recall frequencies*

	First *H*					Second *H*				Total subjects
		C	FS	D	O		C	FS	D	O	
List alpha	Y	6	10	10	2	Y	5	7	2	—	90
	B	2	3	2	—	B	7	7	3	—	
	O	2	14	24	15	O	10	12	12	25	
List beta	Y	5	8	8	—	Y	19	13	3	2	76
	B	—	—	6	—	B	—	4	1	1	
	O	3	19	11	16	O	2	8	11	12	

Y = yellow; B = blue; O = other or unclassified; C = certain; FS = fairly sure; D = doubtful.

and 0·5, and the frequency analysis leads to $\chi^2 = 5{\cdot}35$, $p < 0{\cdot}03$ (two-tail), as shown in Table 3.

This result, taken in conjunction with that reported by Green (1958), flatly contradicts both the gestalt and serial interference predictions. According to Von Restorff, the two *H*

engrams should agglutinate; but we have already shown that succeeding items do not affect the recall of a similar earlier item, and we are left with the fact that under these conditions recall is *facilitated* by a similar preceding item. In previous experiments it has been possible to interpret the results in terms of pro-active interference, but, in this case, the alternative behaviouristic account is equally lost. We are therefore driven to conclude that the relevant variable is structural change.

Table 3. *Comparison of recall frequencies*

		R	F	
Group	Alpha	28	62	90
	Beta	37	39	76
		65	101	166

R = No. subjects recalling colour correctly; F = No. of subjects forgetting.

$\chi^2 = 5{\cdot}35$.

d.f. = 1; $p < 0{\cdot}03$ (two-tail).

Some additional interest attaches to the ratings of subjective conviction. If we draw our comparison between those who were correct *and* certain of it, 6 subjects in the alpha group and 19 in the beta group, instead of all those correct with varying degrees of conviction, the difference between the groups becomes far more distinct. $\chi^2 = 10{\cdot}76$; d.f. = 1; $p < 0{\cdot}005$. A sidelight is also thrown on the question of the reliability of testimony. Of the 61 ratings of certainty only 32 were in fact justified. Nor was this entirely a matter of confusing the first and second letters *H*. Of the 29 'certain' and wrong answers only 12 came under the heading of a simple inversion. Breaking down the 31 ratings of certainty for the critical item we find that for the alpha group, where it was the first letter *H*, the probability of being right when certain is 0·6 (6 out of 10) while for the beta group $p = 0{\cdot}9$ (19 out of 21). So if the event has mild attention-getting value then the testimony would seem to be more reliable ($p = 0{\cdot}06$) two-tail. It would not be wise to assume that the relationship is a simple one. Where the event has attention-getting value because of its high emotional content (such as in the case of a witness to a crime), an entirely different picture might emerge.

The data invite one further observation. Contrary to what has been reported in earlier studies by Blackburn (1945) and Munsterberg (1909), but in agreement with the studies surveyed by Whipple (1909), there appears to be a clear-cut relation between the degree of confidence expressed and the probability of being right. For those who were 'certain' the probability of being correct is 0·525 (32 out of 61). For those who were 'fairly sure' $p = 0{\cdot}286$ (30 out of 105), and for the 'doubtful' replies, $p = 0{\cdot}236$ (22 out of 93). Testing this against the null hypothesis, $\chi^2 = 14{\cdot}6$; d.f. = 2; $p < 0{\cdot}001$. In fact our confidence level is even better when we consider that this value of χ^2 takes no note of the consistent direction of the trend. Blackburn's apparently anomalous result may be simply a function of delay between observing and recalling (his subjects were asked for a report some 2 weeks after the event). Another possibly important variable is the subject's attitude. If the event is staged and the subject does not realize that a report will be required, as with Blackburn's subject, the method of attending may be very different from that adopted in a straightforward memory experiment.

Since the present paper went to press, Newman & Saltz (1958) have confirmed that intra-list interference due to stimulus and response generalization cannot account for isolation effects.

References

Blackburn, J. (1945). *Psychology and the Social Pattern.* London: Kegan Paul.

Gibson, E. J. (1940). A systematic application of the concepts of generalisation and differentiation to verbal learning. *Psychol. Rev.* **47**, 196–229.

Green, R. T. (1956). Surprise as a factor in the von Restorff effect. *J. Exp. Psychol.* **52**, 340–4.

Green, R. T. (1958). Surprise, isolation, and structural change as factors affecting recall in a temporal series. *Brit. J. Psychol.* **49**, 21–30.

Munsterberg, H. (1909). *Psychology and Crime.* London: Fisher and Unwin.

Newman, S. E. & Saltz, E. (1958). Isolation effects: stimulus and response generalization as explanatory concepts. *J. Exp. Psychol.* **55**, 467–72.

Saul, E. V. & Osgood, C. E. (1950). Perceptual organisation of materials as a factor influencing ease of learning and degree of retention. *J. Exp. Psychol.* **40**, 372–9.

Von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld. *Psychol. Forsch.* **18**, 299–342.

Whipple, G. W. (1909). The observer as reporter: a survey of the 'psychology of testimony'. *Psychol. Bull.* **6**, 153–70.

(*Manuscript received* 23 *July* 1957)

THE INFLUENCE OF THE ALBEDO AND COMPLEXITY OF STIMULI ON VISUAL FIXATION IN THE HUMAN INFANT

By D. E. BERLYNE

*Centre for Advanced Study in the Behavioural Sciences**

I. Introduction

For several weeks after birth the human infant divides his time between sleeping, crying and feeding. But, before long, more and more of his time begins to be taken up by a fourth activity, namely, looking. It is clear that his eyes do not wander about the visual field purely at random; his gaze dwells on an object and then shifts to another. Some of the properties that favour fixation of particular objects are fairly obvious and have been reported in the literature—brightness and movement (Hall, 1903; Piaget, 1936), colour (Valentine, 1914; Staples, 1932) and secondary reward value (possessed by human faces and feeding bottles). These variables hardly suffice to account for the direction of all the baby's visual fixations during his first year. But, apart from them, the observing behavour of the human infant seems to have received little study, especially in view of the time and eagerness that he devotes to it. The experiment to be reported tested the effects of two variables, albedo and complexity. Various factors that can be grouped together under the word 'complexity' have turned out to be important determinants of exploratory behaviour in animals (see Berlyne & Slater, 1957) and of orienting responses in human adults (Berlyne, 1957*a*, 1958). They bear interesting relations to variables figuring in information theory and in the theory of conflict (Berlyne, 1957*b*).

II. The experiment

Stimulus material (see Fig. 1). Four series of three rectangular cards were prepared. Series A consisted of plain black, grey, and white rectangles; series B, C and D each contained three figures with different degrees of complexity, but all made up of equal areas of black and white. Card B3 bore a pattern, kindly supplied by Dr F. Attneave (see Attneave, 1954, p. 188), of black and white dots, randomly distributed. Each pattern measured $6 \times 4{\cdot}5$ in. and was surrounded by a $\frac{1}{2}$ in. grey margin. A paired-comparison technique was used: the three cards in each series were presented two at a time side by side, each of the three pairs occurring once in one spatial arrangement and once in the other. There were thus twenty-four trials altogether.

The twenty-four pairs were arranged in a randomized sequence such that each successive set of four pairs contained one from each of the four series of cards. This sequence was used with half of the infants, while the other half had a sequence formed by reversing the temporal order and interchanging left and right in the first sequence.

Subjects. The fourteen subjects ranged in age from 3 to 9 months. Three were adopted babies, while the remainder were children of students and professional people.

Procedure. It was essential for the experimenter who presented the cards to the infants and observed their eye movements not to know which cards were present in any particular trial. Otherwise, there would have been a possibility of introducing constant errors, either in the manner of holding the cards or in the observation of the direction of fixation. For this reason, there were two experimenters who alternated between two roles.

* Now at the University of California, Berkeley. The experiment was carried out and the paper written while the author was on leave of absence from the University of Aberdeen.

The subject was propped up in a sitting position or lay supine, according to his stage of development. For each trial the first experimenter selected the appropriate pair of cards and clipped them side by side and 6 in. apart on a 9×16 in. board. He then handed the board to the second experimenter without allowing him to see the cards. The second experimenter waited until the subject's gaze, spontaneously or by contrivance, was central, whereupon he brought the cards into the infant's field of vision from above, holding the board about 18 in. away. He called out 'right' or 'left' according to the card that the infant *first* fixated, and the first experimenter recorded the result. On one or two occasions, it was impossible to decide which card was fixated first; when this happened, the same pair was presented again later in the sequence.

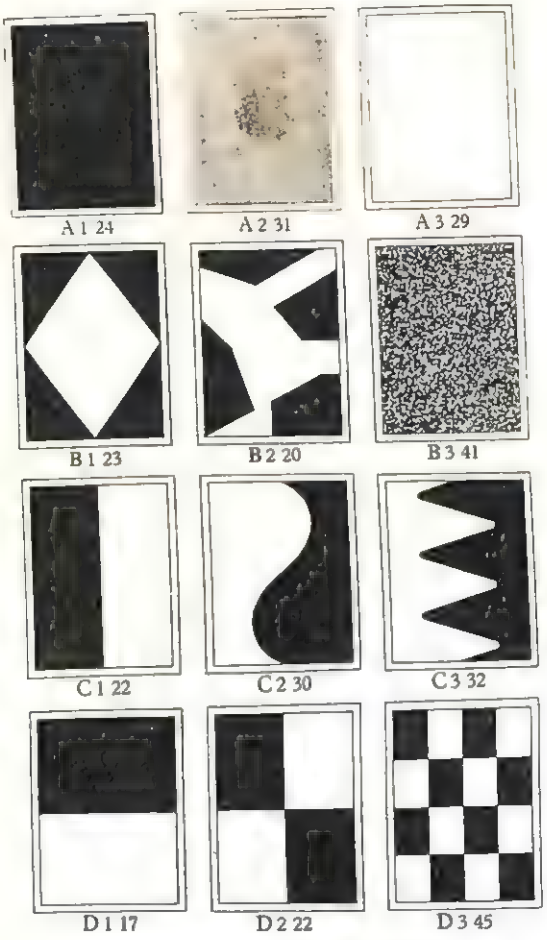


Fig. 1. The stimulus material.

III. Results

The number of times (out of fifty-six presentations) that each pattern was fixated first is shown beneath it in Fig. 1. The results for each series were subjected to two statistical treatments. First, the statistic S (Bradley, 1955, p. 460), which has the χ^2 distribution with 3 D.F., was computed, to ascertain whether there were any significant preferences among the patterns in the series. The relevant data are shown in Table 1, where it will be seen that S is significant for series B and D but not for series A and C. Secondly, the sign test was applied to each paired comparison separately to find out which individual patterns were actually preferred to which. As Table 1 makes clear, patterns B 3 and D 3

Table 1. *First fixations*

Patterns ... Series	1 and 2	1 and 3	2 and 3	S ($=\chi^2$, 3 D.F.)
A	12:16 N.S.	12:16 N.S.	15:13 N.S.	1·29 N.S.
B	16:12 N.S.	7:21 $P<0{\cdot}05$	8:20 $P<0{\cdot}05$	12·71 $P<0{\cdot}01$
C	11:17 N.S.	11:17 N.S.	13:15 N.S.	2·71 N.S.
D	11:17 N.S.	6:22 $P<0{\cdot}01$	5:23 $P<0{\cdot}01$	22·00 $P<0{\cdot}01$

were significantly preferred to other patterns in the same series, but no other preferences emerged. The findings can thus be summarized as follows. (1) Differences in albedo did not affect the direction of the first fixation. (2) Two patterns, B3 and D3, were more likely than others to attract first fixations.

Of these conclusions, the first recalls an unpublished experiment by the writer in which rats were confronted with black and white rectangles and approached both colours equally often. Similarly, Montgomery (1953) found no differences among the amounts of exploratory activity by rats in white, grey and black alleys. Gross differences in intensity of light, on the other hand, appear to influence the responses of rats and the fixations of human infants.

It is noteworthy that the two patterns in the complexity series that were especially attractive were ones with much more *contour* than the others in the same series. A high proportion of the ganglion cells in the vertebrate eye appear to respond either to the onset or to the termination of illumination only (Bartley, 1951, pp. 965–9). Whenever the eyes are in motion, whether in the course of large-scale traverses of the visual field or as a result of the physiological nystagmus described by Marshall & Talbot (1942), the patterns with the largest amounts of contour will produce the most excitation of on- and off-fibres. The high stimulating power that results from this phenomenon might account for the high eye-drawing power of patterns like B3 and D3. The influence of some of the other forms of complexity that have been found to affect curiosity and attention in adults (Berlyne, 1957*a*, 1958) must, on the other hand, be attributed to more central processes.

The author takes pleasure in acknowledging the generous co-operation of Miss D. C. Wertz and Mr T. A. Ludé of the Social Service Division, Department of Public Health and Welfare, San Mateo County, California, who helped in arranging for subjects. He is also greatly indebted to Miss A. Rooke and Mr L. Hommel, who acted as assistants and second experimenters.

References

Attneave, F. (1954) Some informational aspects of visual perception. *Psychol. Rev.* **61**, 183–93.

Bartley, S.H. (1951). The psychophysiology of vision. In Stevens, S. S. (ed.), *Handbook of Experimental Psychology*. New York: Wiley.

Berlyne, D. E. (1957*a*). Conflict and information-theory variables as determinants of human perceptual curiosity. *J. Exp. Psychol.* **53**, 339–404.

Berlyne, D. E. (1957*b*). Conflict and uncertainty: a point of contact between information-theory and behaviour-theory concepts. *Psychol. Rev.* **64**, 329–39.

Berlyne, D. E. (1958). The influence of complexity and change in visual figures on attention. *J. Exp. Psychol.* **55**, 289–96.

BERLYNE, D. E. & SLATER, J. (1957). Perceptual curiosity, exploratory behavour and maze learning. *J. Comp. Physiol. Psychol.* **50**, 230–4.

BRADLEY, R. A. (1955). Rank analysis of incomplete block designs. III. Some large-sample results on estimation and power for a method of paired comparisons. *Biometrika*, **42**, 450–70.

HALL, G. S. (1903). Curiosity and interest. *Ped. Sem.* **10**, 315–58.

MARSHALL, W. H. & TALBOT, S. A. (1942). Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In Klüver, H. (ed.), Visual mechanisms. *Biol. Symp.* **7**. 117–64.

MONTGOMERY, K. C. (1953). Exploratory behaviour as a function of 'similarity' of stimulus situations. *J. Comp. Physiol. Psychol.* **46**, 129–33.

PIAGET, J. (1936). *La naissance de l'intelligence chez l'enfant.* Neuchâtel: Delachaux and Niestlé. (*The Origin of Intelligence in the Child.* London: Routledge and Kegan Paul, 1953.)

STAPLES, R. (1932). The responses of infants to color. *J. Exp. Psychol.* **15**, 119–41.

VALENTINE, C. W. (1914). The colour perception and colour preferences of an infant during its fourth and eighth months. *Brit. J. Psychol.* **6**, 363–86.

(*Manuscript received* 15 *July* 1957)

SKILL AND CHANCE: VARIATIONS IN ESTIMATES OF SKILL WITH AN INCREASING ELEMENT OF CHANCE

By JOHN COHEN, E. J. DEARNALEY and C. E. M. HANSEL

Department of Psychology, University of Manchester

Experiments with three groups of subjects, ranging in age from 13 to 20 years, were carried out to determine how estimates of success in performance are influenced by (*a*) varying the subjective assessment of the degree of skill required to succeed in a task, and (*b*) varying the degree of subjective chance. The findings were as follows. (i) The result of introducing an element of subjective chance into a task requiring some skill is to make the estimates of success in the task progressively reflect the element of subjective chance as it increases in magnitude. (ii) This effect occurs regardless of the degree of skill subjectively felt to be required for the task. (iii) The values of subjective probability of success in performance given by the older subjects are not significantly different from the products of their corresponding values of ψ(skill) and ψ(chance). In the circumstances of these experiments, therefore, the subjective probabilities based upon the subjects' estimates are equivalent to those which the experimenter could independently arrive at if he only had available to him the values of ψ(skill) and ψ(chance) and had multiplied them. (iv) In the case of younger subjects aged 13–14 years, the subjective probabilities of success in performance are somewhat larger than the products of the two separate subjective probabilities. This may be due to a tendency on the part of younger persons to be more confident of success when their estimates relate to a *single* situation involving two sources of uncertainty as compared with when they relate to two *separate* situations, each involving its own source of uncertainty.

I. Introduction

Tasks in general vary from one extreme at which the performer believes the outcome to be almost entirely due to factors outside his control, to the other extreme at which he believes the outcome is almost entirely due to his personal skill, the element of chance being, he believes, negligible. These beliefs do not necessarily correspond with the 'objective' elements of chance in the situation or with the performer's true skill. Thus, some gamblers believe they have devised systems to 'beat' the play of the roulette wheel; or a person may try to explain away his scholastic failure by bemoaning his bad luck. We are here concerned with performances calling for some degree of skill and involving an element of chance which favours or handicaps the performer. The performer may not consciously and explicitly make an estimate of the factor of chance. In principle, however, it is possible to arrange situations so that his estimates of success can be analysed to reveal the relative influence of two component estimates of success. One of these estimates would depend on how skilful he thinks he is when he believes chance is negligible. This we call ψ(skill). The other estimate would depend on how lucky he believes he would be when he recognizes that the outcome depends on chance. This estimate is made when the task is so simple that he has no doubt about his ability to succeed in it when chance is absent. This second estimate we call ψ(chance). The aim of the present experiment is to determine how subjective probabilities of success in performance are affected by variation in these two component estimates. (In a previous publication—Cohen, Dearnaley & Hansel, 1956—we denoted subjective probability of success as ψ_s in contrast to ψ_f which represented subjective probability of failure.)

II. Procedure

The task consisted in throwing steel darts to hit a circular target and remain there. The target was vertically placed on a wall. There were five sets of five darts. The sets included different numbers of darts (0, 1, 2, 3 or 4) which were blunted so that they could not possibly remain on the target however hard they were thrown. The skill required by the task was varied by altering the diameter of the target. There were twelve target sizes ranging from the largest with a diameter of 28 in. to the smallest with a diameter of 1 in. The subject's distance from the target in the experiment was adjusted during a preliminary practice period, the distance being such that he would probably always succeed at the largest target and probably never at the smallest. This distance varied from subject to subject.

Stage (i). In order to obtain a measure of ψ(skill) for each subject in relation to each target, the targets were successively presented to him first in ascending order of size and then in descending order. (An anticipation error was found as would be expected from this procedure.) Each time the target was presented the subject estimated how many times out of five attempts he thought he would succeed when all five darts were sharp. Since it was necessary to present targets yielding the same values of ψ(skill) for all subjects, five target sizes were chosen for each subject at which his mean estimates of success were 1, 2, 3, 4 and 5 times out of 5, respectively, or as near as possible to these values. This provided, for each subject, five targets representing five levels of ψ(skill). Although the actual target sizes varied from subject to subject they had, respectively, the same levels of ψ(skill) for all. We thus obtain a measure of ψ(skill).

Stage (ii). Having established five sizes of target corresponding to five arbitrary chosen values of ψ(skill), when all five darts were sharp and the chance element of drawing a blunt dart was therefore ruled out, we now determined each subject's estimates of success at varying levels of chance, i.e. when there was a possibility that he might draw one or more blunt darts. These estimates were obtained with the largest target at which the subject was sure he would always succeed provided he had five sharp darts. Naturally, before choosing a dart, the subjects could not see which one was blunt and which was sharp. They always knew how many blunt darts there were in each set before choosing a dart. They understood too that the dart would be replaced in the set at random after use.

Stage (iii). The final step was to obtain estimates from each subject when the target size, and hence ψ(skill) and the element of chance, in terms of the number of blunt darts, were varied simultaneously. Thus at each of the remaining four levels of subjective difficulty, estimates were obtained when there were 1, 2, 3 or 4 blunt darts in a set of five darts before each throw. The estimates at this stage and at the preceding one were obtained in random order. This procedure was employed with thirty-one schoolchildren aged 13–14 years and with seven undergraduates.

In a second comparable experiment with eleven undergraduates the task consisted in throwing a bead into a beaker. In order to obtain a measure of subjective difficulty, the beakers presented were of various sizes, thus matching the varying size of the targets of the previous experiments. The subject had to draw a bead from a bowl containing five beads each of which could be either red or green. The varying proportions of red and green beads were known to him, though he was unable to see what was in the bowl. The element of chance was introduced by allowing the subject to throw a bead only if he drew a red one. Thus, in this second experiment ψ(skill) is based upon estimates when the subject is sure that he will always draw a red bead, and ψ(chance) is based upon estimates when the subject is sure he can succeed in throwing a bead into the beaker, but realizes that whether or not he draws a red bead is entirely outside his control.

III. Results

All the estimates in what follows are converted from the number of successes out of five to fractions of unity. These fractions of unity we take to be a measure of subjective probability. The five targets arbitrarily chosen, when the subject knew that all the darts were sharp, were such as to obtain values of ψ(skill) as near as possible to 1·0, 0·8, 0·6, 0·4, and 0·2. The experimental average values of ψ(skill) for these five chosen targets were found to be 1·0, 0·80, 0·61, 0·39 and 0·19. These appear in the first column of Table 1. The corresponding average values of ψ(chance), when the subject knew there was a smaller or larger chance of drawing a blunt dart before each throw were found to be: 1·00, 0·73, 0·54, 0·37 and 0·24. These are shown as column headings in Table 1. They correspond to

a priori probabilities 1·0, 0·8, 0·6, 0·4, 0·2 of drawing a sharp dart and were obtained when the subject was aiming at the largest target, that is, when he thought he would always succeed with five sharp darts.

A comparison of these two latter sets of figures, viz. of ψ(chance) with the *a priori* values show that ψ(chance) is greater than the *a priori* value when this is small and is less than the *a priori* value when it is large. The two sets of figures correspond most closely when the *a priori* value is about 0·3. This supports the outcome repeatedly obtained in earlier experiments (Cohen & Hansel, 1956), which suggests that we tend to overestimate our capacities when we think the task is hard and to underestimate them when we think it is easy, an effect that results in a relative contraction of the subjective probability as compared with the *a priori* scale (Cohen & Hansel, 1957*a*).

Our primary concern here is to determine the variation in estimates of success with variation in the two components, ψ(skill) and ψ(chance). The figures in Table 1 are presented with this in view. In the context of the present experiment ψ(chance) refers to the subjective probability when the probability of drawing a sharp dart varies. This was determined for a target at which the subject thought he would always succeed if he had five sharp darts. Similarly ψ(skill) refers to the subjective probability of hitting the target at any given size when the *a priori* probability of drawing a sharp dart is 1·0 (i.e. when all five darts are sharp).

Table 1. *Mean estimates of success in performance (expressed as subjective probabilities) when both ψ(chance) and ψ(skill) vary*

(N = 31 aged 13–14)

	ψ (chance)†			
ψ(skill)*	0·73	0·54	0·37	0·24
0·80	0·62	0·48	0·33	0·21
0·61	0·56	0·43	0·31	0·24
0·39	0·40	0·29	0·24	0·15
0·19	0·20	0·13	0·12	0·07

* When all five darts are sharp.
† Corresponding to *a priori* probabilities 0·8, 0·6, 0·4 and 0·2 of drawing a sharp dart.

It would be reasonable to expect estimates of success, at any given size of target, to decline as the chance of choosing a sharp dart diminishes. In other words, at any given value of ψ(skill), estimates of success would decline with decrease in ψ(chance) and vice versa. This was, in fact, found to be the case.

The question, however, remains as to the magnitude of decline to be expected. On common sense grounds we should expect the decline progressively to reflect the diminishing values of ψ(chance). If subjective probabilities combined in a manner analogous to mathematical probabilities, their magnitude would be approximately the product of corresponding values of ψ(skill) and ψ(chance). We have computed these products and subtracted them from the corresponding observed values in Table 1. The resulting residuals are presented in Table 2. They are consistently positive. The mean discrepancy between the observed and theoretical values is 0·06 which, contrary to the null hypothesis, is significant at the 0·001 level (t = 7·62, D.F. = 15, $P < 0{\cdot}001$). The discrepancy is significant but small. We may conclude therefore that the ψ-values in Table 1 are not entirely reducible to the product of two separate subjective probabilities.

Table 1*a*. *Table of residuals when products of ψ(chance) and ψ(skill) are subtracted from the corresponding observed values*

0·04	0·05	0·03	0·02
0·11	0·10	0·08	0·09
0·12	0·08	0·10	0·06
0·06	0·03	0·05	0·02

When we turn, however, to the results of the similar experiment with the undergraduates we do not find any significant difference between the calculated and observed values. The mean of the residuals is zero. The parallel data for these older subjects are presented in Tables 2 and 2*a*.

Table 2. *Mean estimates of success in performance (expressed as subjective probabilities) when both ψ(chance) and ψ(skill) vary*

	ψ(chance)†			
ψ(skill)*	0·69	0·46	0·38	0·14
0·83	0·46	0·37	0·26	0·17
0·59	0·49	0·31	0·23	0·11
0·40	0·29	0·20	0·09	0·06
0·15	0·20	0·03	0·03	0·00

* When all five darts are sharp.
† Corresponding to *a priori* probabilities 0·8, 0·6, 0·4 and 0·2 of drawing a sharp dart.

Table 2*a*. *Table of residuals when products of ψ(chance) and ψ(skill) are subtracted from the corresponding observed values*

−0·11	−0·01	−0·06	+0·06
+0·08	+0·04	−0·01	+0·03
+0·01	−0·02	−0·06	0·00
+0·01	−0·04	−0·03	−0·02

The same agreement between calculated and observed values in the case of the older subjects appeared in the second experiment in which drawing and throwing a bead replaced selecting and aiming a dart. The relevant figures are shown in Tables 3 and 3*a*.

Table 3. *Mean estimates of success in performance (expressed as subjective probabilities) when both ψ(chance) and ψ(skill) vary*

	ψ(chance)†			
ψ(skill)*	0·69	0·53	0·35	0·18
0·81	0·52	0·40	0·30	0·16
0·60	0·50	0·34	0·22	0·08
0·43	0·23	0·22	0·13	0·02
0·21	0·16	0·09	0·00	0·00

* When all five beads are red.
† Corresponding to *a priori* probabilities 0·8, 0·6, 0·4 and 0·2 of drawing a red bead.

Table 3*a*. *Table of residuals when products of ψ(chance) and ψ(skill) are subtracted from the corresponding observed values*

−0·4	−0·3	+0·2	+0·1
+0·9	+0·2	+0·1	−0·3
−0·7	−0·1	−0·2	−0·6
+0·2	−0·2	−0·7	−0·4

In venturing to interpret these results we may point to a greater internal consistency in the judgements of the older subjects as compared with those of the younger ones. The judgements of the older subjects in Stage (iii) logically follow from those they made at Stages (i) and (ii). In the case of the younger subjects, however, there is some departure from this logical consistency. The departure may be due to a tendency on the part of the younger subjects, when faced with two sources of uncertainty, to be relatively less affected by it when both sources are present in the same situation than when they occur in two separate situations. More specifically, estimates of their skill made by the younger subjects cannot be divorced in their minds from their feelings about the chances of drawing a sharp dart or not, and vice versa. The older subjects display similar tendencies in the estimates they make in the two different experimental situations, with darts and beads, respectively. Indeed, their estimates may be expressed as a compound of two component subjective probabilities, in a manner analogous to the combination of mathematical probabilities.

The fact that similar results were obtained with two different tasks and different methods of introducing an element of 'subjective chance' indicates that the effects are not specific to a particular situation but appear to have a degree of generality. Hence, the outcome of the experiments with the older subjects is that in a complex situation involving two sources of uncertainty, the subjective probability of success in performance is equivalent to the product of the ψ-values relating to each separate source. Any possibility that the subjects (younger or older) might calculate their estimates in terms of the two components was completely ruled out by the randomizing procedure in the experiment. In any case, the younger subjects could certainly not have known that they could calculate their estimates of success by multiplying two component estimates.

These results suggest a number of problems that need to be investigated, especially in the light of our previous experiments (Cohen & Hansel, 1957*b*) in which we studied the mode of combination of values of ψ(chance) deriving from separate sources of uncertainty. In further experiments we have examined estimates made when (*a*) there are two or more *a priori* sources of uncertainty simultaneously present, and (*b*) when there are two or more sources of uncertainty about skill simultaneously present. The results of these subsequent experiments are reported elsewhere (Cohen, Dearnaley & Hansel, 1958).

REFERENCES

COHEN, J. & HANSEL, C. E. M. (1956). *Risk and Gambling.* London: Longmans Green.

COHEN, J. & HANSEL, C. E. M. (1957*a*). La répartition des probabilités subjectives. *J. Psychol.* January–March, 10–21.

COHEN, J. & HANSEL, C. E. M. (1957*b*). The nature of decisions in gambling. *Acta Psychol.* **13**, 357–70.

COHEN, J., DEARNALEY, E. J. & HANSEL, C. E. M. (1956). The addition of subjective probabilities. *Acta Psychol.* **12**, 371–80.

COHEN, J., DEARNLEY, E. J. & HANSEL, C. E. M. (1958). *Durham Research Review* (in the Press).

(*Manuscript received* 10 *July* 1957)

EFFECT OF QUINAL BARBITONE, DEXTRO-AMPHETAMINE, AND PLACEBO ON APPARENT TIME*

By SANFORD GOLDSTONE, WILLIAM K. BOARDMAN and WILLIAM T. LHAMON

Department of Psychiatry, Baylor University College of Medicine, Houston, Texas

Quinal barbitone, dextro-amphetamine and placebo were compared with respect to changes produced in perceived duration in ninety normal subjects. A method was employed which permitted quantitative derivation of a subject's concept of *one second*, pre-drug, 30 min. post-drug and 60 min. post-drug. The following findings were obtained. (1) A significant decrease in the clock-measured value of the apparent second with the ingestion of dextro-amphetamine. (2) A significant increase in the clock-measured value of the apparent second with the ingestion of quinal barbitone. (3) A trend toward increase in the apparent second with placebo. These findings were discussed in terms of earlier reports and a theoretical framework was offered.

I. Introduction

Although alterations in the estimation of time are among the most frequently reported perceptual changes associated with drug states, little quantitative evidence is available which either verifies or refutes published and accepted notions based upon qualitative, observational case study. The lack of experimentally defined alteration in the time sense has led to considerable confusion with respect to the meaning of such words as underestimation and overestimation, physical time and personal time. This paper describes a method of studying the effects of common drugs upon apparent time, where prior qualitative, observational evidence was available to suggest that these agents might influence perceived duration. Sedation is frequently accompanied by the subjective experience of external time 'flying'—the standard physical clock dimension is underestimated. Stimulation, on the other hand, is more frequently associated with the subjective experience of external time 'dragging'—clock time is overestimated.

A method was employed which permitted a quantitative derivation of the concept of one clock-second or Second Estimation Point (S.E.P.). Considerable data on normal (Goldstone, Lhamon & Boardman, 1957; Smythe & Goldstone, 1957) subjects was available, and the method was sensitive to the reported time-estimation changes associated with schizophrenia (Lhamon & Goldstone, 1956; Weinstein, 1956) and lysergic acid diethylamide (Boardman, Goldstone & Lhamon, 1957). It was hypothesized that a drug-induced awareness of apparent slowing of clock time (i.e. time 'dragging') would accompany a smaller S.E.P. A drug-induced awareness of an apparent speeding of clock time (i.e. time 'flying') would accompany a longer S.E.P. in the direction of underestimation.

Drug-induced alterations in apparent duration have been reported with marihuana (Bromberg, 1934), lysergic acid and mescaline (Hoch, Cattel & Pennes, 1952), and nitrous oxide (Steinberg, 1955). Only in the nitrous-oxide study was an experimental, quantitative

* This research was supported in part by Grant M-1121, United States Public Health Service and the United States Air Force Office of Scientific Research, Contract AF 18 (603)-79 monitored by the Office of Scientific Research of the Air Research and Development Command.

procedure with a group of subjects reported. Steinberg (1955) investigated the differences in estimation of 30 sec. by counting aloud at an estimated speed of one count per second between an anaesthetized and a control group. Both groups demonstrated a significant increase in the apparent 30 sec. after drug with no difference between the two groups. She concluded as follows: 'The drug therefore cannot be regarded as having had any effect on estimating seconds. That the mean scores of both groups should have risen at all may possibly be an artifact attributable to the apparatus used to administer the gas. . . .' The anaesthetized group did demonstrate a change in time estimation following the gas, but the author preferred to dismiss her finding, presumably in favour of the assumption that a placebo is an inert agent with little potential for altering behaviour. The present study reduced the cues available in the common counting procedure and employed a stimulant in addition to the sedative and placebo, assuming that placebos are not always inert and may produce psychological effects which when similar to drug effects may tend to mask the effect of the active agent.

II. METHOD AND PROCEDURE

Apparatus. An electronic timer activated an audio oscillator that delivered, through headphones, a tone of 725 cyc./sec. for durations of between 0·1 and 2·0 sec. at 0·1 sec. step intervals. The amplitude of the tones was comfortable and constant.

Subjects. Ninety medical students and interns were divided into three equal groups, each receiving one of the studied drugs. One group received 0·2 g. quinal barbitone (called in America secobarbital), a second group received 15 mg. dextro-amphetamine, and a third group received a lactose placebo. All drugs were administered orally.

Procedure. Subjects were studied under three conditions: (1) pre-drug, (2) 30 min. post-drug, (3) 60 min. post-drug. They were presented with a series of audible durations in accordance with a method of limits technique and were required to respond to each duration with the report 'more' or 'less', indicating a judgement of more or less than 1·0 sec. Twenty alternating runs of ascending and descending durations were provided starting at 1·0 sec., with 0·1 a sec. step interval. If the subject reported *more* to the initial 1·0 sec. stimulus, the first run was of a descending nature and terminated when he reported *less* three consecutive times. If the subject reported *less* to the initial 1·0 sec. stimulus, the first run was of an ascending nature and terminated when he reported *more* three consecutive times. This procedure was followed for each drug under the pre-drug, 30 and 60 min. post-drug conditions.

The percentages of reports *less* for each duration of each drug condition were plotted on normal probability paper on the assumption that these values were normally distributed and the straight line visually fitted (Dixon & Massey, 1951). The S.E.P. or duration at which the subject reported more and less 50% of the time were derived from these plots.

III. RESULTS

All S.E.P.'s were found to be log normally distributed and the appropriate transformations were made. The obtained pre-drug mean log S.E.P.'s for the quinal barbitone, dextro-amphetamine, and placebo groups of $\bar{1}\cdot7691$, $\bar{1}\cdot8380$ and $\bar{1}\cdot8132$, respectively, did not differ significantly from each other. The antilogs of the mean log S.E.P.'s (geometric mean S.E.P.'s) of these three pre-drug groups (0·59, 0·69 and 0·65 sec., respectively) indicate that all groups tended to overestimate the value of the clock second prior to drug administration. These results correspond to previous findings (Goldstone *et al.* 1957). The differences between the pre-drug log S.E.P., and 30 and 60 min. post-drug log S.E.P.'s were obtained. Table 1 summarizes the mean shift in log S.E.P. for each drug from the pre-drug to the 30 and 60 min. post-drug conditions. A rise in log S.E.P. from pre-drug to a post-drug condition is designated by a *plus* sign and a reduction in log S.E.P. appears with a *minus* sign. Mean log S.E.P. change is also interpreted in terms of percentage change for convenience and clarity. Change in log S.E.P. was examined with a *t*-test for paired observations.

Table 1. *Mean shift in log* S.E.P. *for each drug group from pre-drug to* 30 *and* 60 *min. post-drug in log units and in percentage* (*a* plus *sign reflects a rise in* S.E.P. *and a* minus *sign reflects a reduction in* S.E.P.)

	Quinal barbitone		Dextro-amphetamine		Placebo	
	Log shift	% shift	Log shift	% shift	Log shift	% shift
30 min. post-drug	+0·0475	+11·9	−0·0430	−10·3	+0·0294	+7·1
60 min. post-drug	+0·0876	+22·2	−0·0597	−15·1	+0·0363	+8·7

The quinal barbitone group demonstrated a significant rise in log S.E.P. after 60 min. ($t=2·83$; $P<0·01$) with a smaller rise noted after 30 min. ($t=1·81$; $P<0·10$). This alteration in apparent duration is in the direction of underestimation and is compatible with a subjective report that time seems to be 'flying'. The dextro-amphetamine group demonstrated a significant reduction in log S.E.P. after both 30 min. ($t=2·04$; $P<0·05$) and 60 min. ($t=2·60$; $P<0·025$). This alteration in experienced duration is in the direction of overestimation and is compatible with a subjective report that time seems to be 'dragging'. The placebo group demonstrated a small but not significant rise in log S.E.P. after both 30 and 60 min. The similar direction of both placebo and sedative groups conforms to the previously reported finding of Steinberg (1955).

The quinal barbitone group differed significantly from the dextro-amphetamine group with respect to change in log S.E.P. after both 30 min. ($t=2·70$; $P<0·01$) and 60 min. ($t=3·84$; $P<0·005$). The placebo group differed significantly from the dextro-amphetamine group with respect to change in log S.E.P. after both 30 min. ($t=2·92$; $P<0·01$) and 60 min. ($t=2·55$; $P<0·025$). No significant differences were obtained between placebo and quinal barbitone after 30 or 60 min., quinal barbitone after 30 min. and quinal barbitone after 60 min., dextro-amphetamine after 30 min. and dextro-amphetamine after 60 min., or placebo after 30 min. and placebo after 60 min. with respect to log S.E.P. change. It is of interest that the increase or decrease in log S.E.P. for each drug was greater after 60 min. than after 30 min.

These results demonstrate a significant decrease in the clock-measured value of the apparent second with dextro-amphetamine; a significant increase in the clock-measured value of the apparent second with quinal barbitone, and a trend toward increase in the apparent second with placebo. The effect of placebo and quinal barbitone upon the apparent second differ significantly from the effect of dextro-amphetamine.

Subjects were asked to report the effects of the drugs and to guess the nature of the medication. Most of the quinal barbitone group reported sedation at 30 min. and all but one experienced the usual symptoms at 60 min. Several subjects had to be awakened to be tested and two fell asleep during testing. None of the dextro-amphetamine group reported stimulation of mental or physiologic processes while five reported drowsiness at 30 min. One subject guessed that he had a stimulant, while three continued to feel drowsy at 60 min. Most of the dextro-amphetamine subjects reported no noticeable drug effects. This suggests that time-sense measurements provide a sensitive method for demonstrating changed mental functioning, even prior to subjective awareness. Ten of the placebo group thought they had a mild sedative and one subject fell asleep.

Subjects were also asked to report their feeling about how consistent they were from test to test. Most reported experiencing the same duration as a second, even though the

clock indicated a marked change. The experienced second had apparently remained the same, providing evidence for the apparent speeded or slowed perception of time. The drug-induced change in the clock-measured second prevailed even when subjects reported continued use of such mental devices as counting or visualizing clock movement.

IV. DISCUSSION

These results quantitatively demonstrate the common qualitative observation of a drug-induced alteration in the estimation of time. The experience of time flying might be interpreted as a relative speeding of the sequence of external events when compared to a personal standard, or frame of reference; while the experience of time dragging might be interpreted as a relative slowing of the sequence of external events when compared to one's standard. Pieron (1952) suspected that appreciation of time was associated with a subjective standard frame of reference or 'time appropriate to our organism'. He thought that changes of internal temperature must modify this standard. François (1928) found that the apparent second was shortened when body temperature rose. Hoagland (1933) verified this finding and hypothesized that increased body temperature should be associated with a speeded internal chemical clock that should correlate with the perception of external time proceeding more slowly, since more 'chemical clock time' would pass in a given physical clock unit than under normal conditions. Pieron (1952) reports that when it is warmer trained bees fetch sugar ahead of time, and when it is colder they are late.

Without committing ourselves to the source and nature of the internal pacemaker, we might make use of Hoagland's and Pieron's formulations about an internal temporal standard. Let us postulate the organism as a clock and organismic time as a function of an event-counting process which is independent of arbitrary physical clock units. These 'events' are hypothetical constructs which at present are not identified in terms of any specific biologic, chemical or psychologic system, but represent discrete recurrences, the counting of which provide us units of personal time. It is further postulated that each organism has a standard rate of countable events which varies, within limits, with modifications of organismic or environmental conditions. It is this variation in event rate which we associate with alteration in perceived time. An earlier study (Smythe & Goldstone, 1957) suggested that at age 8 the learning of culturally determined time standards takes place. Since individual differences probably exist in event rate and personal time, cultural standards permit people to conform their private rates with consensually accepted physical time and allow communication and interpersonal transactions with respect to duration, sequence and order. We suspect that at age 8 a standard and stable event rate becomes identified with newly learned physical time concepts. The speeding or slowing of event rate would modify the correspondence with physical units and alter the physical clock value of the apparent unit. If a rise in body temperature is associated with an acceleration of countable events, more events will transpire in a given unit of clock time. If the given unit (let us say 1 sec.) had formerly been identified with a specific number of events (let us say ten), this number will pass in less time (let us say 0·5 sec.) providing an overestimation of the clock second (in this case, an apparent second or S.E.P. of 0·5 sec.). This formulation corresponds to the obtained findings with raised temperature.

The stimulating effect of dextro-amphetamine also leads to a prediction of a speeded event rate which should be associated with a shorter clock-measured apparent second and perception of 'things slowed down'. On the other hand, the sedative effects of quinal barbitone might lead to a prediction of a slowed event rate which should be associated with a larger clock-measured apparent second and the perception of 'things speeded up'. These predictions were verified by the results of this study. Other reports (Rosenthal & Frank, 1956) have indicated placebo effects that are similar to those of a sedative. It is possible that the small placebo effect and the larger barbiturate effect in the same direction might merely be a function of time passage or the repetition of the test. However, the magnitude and consistency of the barbiturate effect suggests that the drug rather than test conditions was the source of judgement alteration. Subjects studied throughout a single day and regularly for several months (Boardman *et al.* 1957) without a drug did not demonstrate the significant rise in S.E.P. noted with the quinal barbitone group.

The findings of this study conform to those obtained by Steinberg (1955), suggesting that the increased apparent second noted in both the nitrous-oxide and placebo-gas groups was not spurious. It is suggested that a placebo acts much like a sedative in altering perceived durations. These results and the proposed theoretical speculations are in line with Schilder's (1942) description of Sterzinger's findings on retarded and speeded metabolic processes with quinine and thyroxin. Thyroxin administration was associated with relative overestimation of time, while quinine ingestion was associated with a relative underestimation of time.

References

Boardman, W. K., Goldstone, S. & Lhamon, W. T. (1957). The effects of LSD on the time sense of normals: a preliminary report. *Arch. Neurol. Psychiat.* **78**, 321.

Bromberg, W. (1934). Marihuana intoxication. *Amer. J. Psychiat.* **14**, part 1, 303–30.

Dixon, W. J. & Massey, F. (1951). *Introduction to Statistical Analysis.* New York: McGraw-Hill.

François, M. (1928). Contribution à l'étude du sens du temps. *Année psychol.* **28**, 186.

Hoagland, H. (1933). The physiological control of judgements of duration: evidence for a chemical clock. *J. Gen. Psychol.* **9**, 267.

Hoch, P. H., Cattell, J. P. & Pennes, H. H. (1952). Effects of mescaline and lysergic acid (*d*-LSD-25). *Amer. J. Psychiat.* **108**, 579.

Lhamon, W. T. & Goldstone, S. (1956). The time sense: estimation of one second durations by schizophrenic patients. *Arch. Neurol. Psychiat.* **76**, 625.

Goldstone, S., Lhamon, W. T. & Boardman, W. K. (1957). The time sense: anchor effects and apparent duration. *J. Psychol.* **44**, 145.

Pieron, H. (1952). *The Sensations: Their Functions, Processes and Mechanisms*, translated by M. H. Pirenne and B. C. Abbott. New Haven, Conn: Yale University Press.

Rosenthal, D. & Frank, J. D. (1956). Psychotherapy and the placebo effect. *Psychol. Bull.* **53**, 294.

Schilder, P. (1942). *Mind: Perception and Thought in their Constructive Aspects.* New York: Columbia University Press.

Smythe, E. J. & Goldstone, S. (1957). The time sense: a normative genetic study of the development of time perception. *Percept. and Motor Skills*, **7**, 49.

Steinberg, H. (1955). Changes in time perception induced by an anaesthetic drug. *Brit. J. Psychol.* **46**, 273.

Weinstein, A. D. (1956). The differential effects of anchors and anchor sequences on the temporal judgments of short intervals by schizophrenic and normal subjects. Unpublished doctoral dissertation, University of Houston.

(*Manuscript received* 1 *July* 1957)

EXTREME RESPONSE SETS AS A MEASURE OF INTOLERANCE OF AMBIGUITY

By M. I. SOUEIF

Institute of Psychiatry, Maudsley Hospital, London

A *Personal Friends Questionnaire* was administered to various groups of Egyptian subjects totalling 1028 in all. This questionnaire is composed of seventy items, each of which is to be checked either +2, +1, 0, −1 or −2. Extreme response scores were obtained by counting the number of times +2 or −2 were assigned by each subject. Corrected split-half reliabilities of the extreme response scores for males and females were 0·92 in both cases. The extreme response set scores are suggested as possible measures of intolerance of ambiguity. A general hypothesis was formulated according to which a social group with a higher tension level would earn a higher extreme response score than a social group with a lower tension level. The theoretical formulation of the problem in this way allowed for four predictions, about the effect on extreme response set scores of age, sex, membership of a religious minority and socio-economic status. For verification of the predictions, the social groups were matched for age, sex and religion. The four expectations were, on the whole, fulfilled.

I. Introduction

Cronbach (1950) has observed that response sets become more influential as test items become difficult or ambiguous. The crucial point, according to him, is to determine whether these response sets 'are characteristics of the individual stable from time to time, or are transient sets which can only be regarded as errors in testing rather than personality characteristics'. Berg (1953) has demonstrated that 'the tendency to choose extreme positions when marking choices on the *perceptual reaction test* (P.R.T.) and the *word reaction test* (W.R.T.) is acceptably stable over test-retest periods up to 15 days'. Rundquist (1950) reported a correlation of 0·40 ($N = 111$) between this tendency 'in two situations, one immediately following the other'. As Berg (1953) puts it, 'the indications are that at least some forms of response sets are reflexions of personality traits'. Some evidence that this is the case appears in scattered studies. Lewis & Taylor (1955) have demonstrated the tendency for high anxious subjects (who scored twenty-five or above on the Taylor Anxiety Scale) to have higher extreme responses (*like very much, dislike very much*) on the P.R.T. than the low anxious subjects (scoring seven or below). Berg & Collier (1953) have reported significant differences between the mean extreme responses of high anxious males, high anxious females and generally neurotic males, and low anxious males, low anxious females and normal males, respectively. They demonstrated also that White females and Negro males had significantly higher extreme responses than White males. It might be interesting to note that Terman & Miles (1936, p. 38) called attention 'to a persistent sex difference in the frequency with which male and female subjects check extreme responses on tests of emotional and ethical attitudes'.

The nature of the extreme response sets seems to point to the feasibility of considering the extreme response set scores as a measure of *intolerance of ambiguity*. In reporting findings from her extensive researches, Frenkel-Brunswik (1951, p. 394) has summarized her general results as follows. 'We found evidence, in certain groups of our subjects, of intolerance of ambiguity in the perceptual as well as in the emotional and social areas.' In these groups of subjects 'feelings are expressed in terms of the ends of a continuum

rather than of a continuum proper'. Frenkel-Brunswik further emphasized that 'underlying anxiety issuing from the confusion of one's social identity and from other conflicts is apparently so great that it hampers individuals (in such groups) in facing' external ambiguity (p. 395).

The main objective of this investigation was to study the influence of being a member of a certain social group on intolerance of ambiguity, as measured by extreme response set scores. A general hypothesis was formulated as follows. *Social groups, with different levels of tension, differ from each other with regard to their average intolerance of ambiguity, as measured by number of extreme scores. Other conditions being equal, a social group assumed to have a higher tension level (due mainly to feelings of insecurity) would tend to make more extreme responses when compared with another social group assumed to have a lower tension level.* The concept of *tension* has been widely used in modern psychological literature (Eysenck, 1953, pp. 299–303). In this investigation, by ascribing a level of tension to social groups, a certain average tension level was assumed for the members of each of those groups. With this general hypothesis in mind the following predictions were formed. (*a*) Adolescents are expected to produce more extreme responses than adults. (*b*) Christians (as members of a religious minority group in Egyptian society) are expected to have more extreme responses than Moslems. (*c*) Females are expected to have more extreme responses than males. (*d*) Members of the lower middle class are expected to produce more extreme responses than members of the upper middle class.

II. Procedure

A questionnaire on the characteristics of personal friends, hereafter referred to as P.F.Q., was constructed (see Winslow & Frankel, 1941). It was composed of seventy items representing all the views expressed by twenty Egyptian subjects (fifteen males and five females, ranging in age from 16 to 28) in essays on friendship. Most of the items were single words, e.g. co-operative, dependable, gloomy, ignorant. The items followed each other randomly, without any predetermined arrangement. The respondent was to consider whether in making friends with a given person he (or she) considered each characteristic in the list as *very necessary, desirable, indifferent, undesirable,* or *definitely opposed.* Numerical values assigned to these various responses ranged from +2 through +1, 0, −1 to −2. The P.F.Q. was presumed to be fairly 'unstructured or ambiguous in the sense that there were no "right" or "wrong" answers' (Berg, 1953). Finally, each test was scored for extreme responses by counting the number of times an extreme response (either +2 or −2) was assigned.

To determine the reliability of extreme response sets, a split-half method (i.e. splitting the whole seventy items into odds and evens) was adopted. Estimated reliability for males was 0·85 ($N = 100$). Corrected for length by the Spearman-Brown formula it turned out to be 0·92. The corresponding values for females were 0·86 and 0·92, respectively ($N = 100$).

An attempt was made to verify the assumption that extreme response scores were not affected by the fact that some of the subjects signed their responses. The P.F.Q. was administered to 338 adolescent females (Moslems). The accompanying instructions were that they were free to sign their responses or to leave them unsigned. Responses were signed by 174 subjects, while 164 left them unsigned. The medians of both groups were compared, and the difference was not significant. (The distribution of scores in the non-signing group was significantly positively skewed, $t = 2{\cdot}24$.)

The P.F.Q. was administered (in its Arabic version) to various groups of Egyptian subjects totalling 1028 in all. They ranged in age from 12 to 46 with a median of approximately 17 years. Most of the subjects were members of the urban middle class with its lower and upper strata. They were drawn from students of Cairo University, from Egyptian secondary school classes in Cairo, Alexandria and some other big towns nearer to the rural districts, and from some Egyptian clubs, during the years 1951–2. The P.F.Q. was usually administered to groups of forty to fifty subjects at a time. The questionnaire sheets had the necessary instructions printed at the top. The subjects were always told that they were free to sign the questionnaire or to leave it unsigned.

III. Results and discussion

To control the effect of certain variables which might have interfered and distorted the results, it appeared desirable to match the groups to be compared, for age, sex and religion. Inspection of Tables 1 and 2 indicates that on the whole the predictions were verified. Thus, for the first prediction, the median score of extreme responses of the

Table 1. *Extreme response scores: means and medians*†

Group	*N*	Mean	S.D.	Median
Adolescents, females, Moslems	338	30·58	9·94	30·12
Adolescents, females, Christians	52	36·42	11·99	34·50
Adolescents, males, Moslems	366	33·38	10·91	33·05
Adolescents, males, Christians	31	32·32	10·94	—
Early adolescents, females, Moslems	54	31·18	9·28	—
Early adolescents, females, Christians	10	42·30	13·64	—
Early adolescents, males, Moslems	235	33·83	10·61	—
Early adolescents, males, Christians	17	34·88	10·39	—
Late adolescents, females, Moslems	284	30·50	9·91	—
Late adolescents, females, Christians	42	35·07	11·11	—
Late adolescents, males, Moslems	131	32·39	11·43	—
Late adolescents, males, Christians	14	29·14	9·67	—
Adults, females, Moslems	67	30·88	9·50	29·30
Adults, females, Christians	25	36·00	10·77	—
Adults, males, Moslems	114	27·31	8·98	25·57
Adults, males, Christians	35	27·28	8·23	—
Upper middle class (males, adolescents, Moslems)	23	31·30	10·41	—
Lower middle class (males, adolescents, Moslems)	21	48·10	10·80	—
Signing group (adolescent females, Moslems)	174	30·70	9·56	30·37
Non-signing group (adolescent females, Moslems)	164	30·45	10·32	29·69

† The medians are stated where they are needed for statistical comparisons owing to skewness of distributions.

Table 2. *Level of significance of differences in incidence of extreme responses:*† *means and medians*

Groups compared	Statistics compared	*t*
Adolescents, males, Moslems—adults, males, Moslems	Medians	5·91***
Adolescents, males, Christians—adults, males, Christians	Means	2·06**
Adolescents, females, Moslems—adults, females, Moslems	Medians	0·51
Adolescents, females, Christians—adults, females, Christians	Means	0·15
Christians, males, adults—Moslems, males, adults	Medians	0·92
Christians, males, adolescents—Moslems, males, adolescents	Means	0·42
Christians, females, adults—Moslems, females, adults	Means	2·09**
Christians, females, adolescents—Moslems, females, adolescents	Medians	2·00**
Females, adults, Moslems—males, adults, Moslems	Medians	2·09**
Females, adolescents, Moslems—males, adolescents, Moslems	Medians	3·06**
Females, adults, Christians—males, adults, Christians	Means	3·33***
Females, adolescents, Christians—males, adolescents, Christians	Means	1·58
Females, early adolescents, Moslems—males, early adolescents, Moslems	Means	1·84
Females, late adolescents, Moslems—males, late adolescents, Moslems	Means	1·63
Females, early adolescents, Christians—males, early adolescents, Christians	Means	1·43
Females, late adolescents, Christians—males, late adolescents, Christians	Means	1·91*
Lower middle class, adolescent males, Moslems—upper middle class, adolescent males, Moslems	Means	5·12***
Signing group, females, adolescent Moslems—non-signing group, females, adolescent Moslems	Medians	0·50

† One tail test in the cases where differences agreed with predictions.
* Significant at the 5% level.
** Significant at the 2% level.
*** Significant beyond 0·001 level.

adolescents turned out to be significantly higher than the median score of adults in the case of Moslem males (beyond 0·001 level). (The medians were used for comparison whenever the distribution of scores in at least one of the two compared groups was significantly skewed.) In the case of Christian males the mean score of adolescents was also significantly higher than the mean score of adults (beyond 0·02 level, using one tail test). In the case of Moslem females the difference between the medians was not statistically significant though it was in the predicted direction. That also was the case with the Christian females. The second prediction was supported in the case of adult and adolescent females. Thus, the mean score of Christian female adults and the median of Christian female adolescents were both significantly higher than the corresponding mean and median of Moslem female adults and adolescents, respectively (both significant beyond 0·02 level). Comparing the scores of Christian adult males and Moslem adult males it was found that the median of the former was higher than that of the latter. This difference, however, did not reach an acceptable level of significance. In the case of adolescent males the difference between Christian and Moslem means was contrary to prediction, though not significant.

The third prediction was supported in the case of adult Moslems and Christians. In those two groups the median of the females was significantly higher than the median of the males. In the case of the Moslem group the difference was significant beyond the 0·02 level, and in the case of the Christian group it was significant beyond the 0·001 level.

The difference between the medians in the Moslem adolescents group turned out to be statistically significant in a direction opposite to that predicted. Thus, the median of the males was significantly higher than the median of the females. This unexpected result warranted a closer look at the distributions of the groups compared. With regard to chronological age the two groups could not be considered homogeneous. Whereas 64% of the males fell into the early adolescence period (under 16), only 16% of the females fell into this category. It was felt desirable to break down each of the two groups into its early and late component subgroups. Comparing the early and the late female subgroups to the corresponding early and late male subgroups did not change the direction of the results, but rendered them not statistically significant.

In the case of Christian adolescents the mean score of the females was higher than the mean score of the males, although the difference was not significant. When comparing the mean score of the late adolescent females with the mean score of the late males, the former was significantly higher (at 0·05 level).

The specific method adopted to verify the fourth prediction was necessitated by certain facts of the social situation. From among the sample of Moslem male adolescents two groups of twenty-three and twenty-one subjects were selected for comparison. The two groups belonged to two secondary schools which differed from each other with regard to the socio-economic background of their populations. One of the two schools was situated in a district inhabited by well-to-do families, and received the majority of its students from among upper middle-class families. The other was situated in a native district, mostly inhabited by lower middle-class families, and received its students largely from among lower middle-class families. Moreover, inspection of the fathers' occupations in the two groups seemed to point to the feasibility of considering them as belonging to those two different socio-economic classes. The two groups were matched for age. The mean age for the upper group was 14·3 ($\pm$0·86) and for the lower group 13·9 ($\pm$1·02). Comparing

both for extreme responses the lower group earned a strikingly higher mean score than the upper group (significant beyond 0·001 level).

In general, the results of this investigation confirm the hypothesis that a social group with a higher tension level tends to manifest higher intolerance of ambiguity, in terms of extreme responses, than a social group with a lower level of tension. The results are, on the whole, in line with and further illuminate the results established by Berg & Collier (1953). Experimental control of age and socio-economic status could possibly have helped in establishing more consistent findings in their case. The results of this investigation, however, seemed to deviate from the theoretical expectation in one respect: Instead of finding that female Moslem adolescents made significantly higher extreme responses than the male Moslem adolescents, the opposite was found to be the case. This unexpected result might be due to the interference of the socio-economic variable which could not be controlled. A general inspection of the fathers' occupations in both samples tended to verify this hypothesis, which, however, needs further investigation.

The theory elaborated in this investigation is consistent with several theoretical suggestions introduced by Frenkel-Brunswik (1949, 1951) through her clinical and experimental researches. In reporting findings about children who were markedly intolerant for perceptual ambiguity, she emphasized that: 'Data on the parents of the children in the rigid, intolerant group reveal that it is their feeling of social and economic marginality in relation to the group to which they aspire from which ensues the desperate clinging to external and rigid rules...' (1949). She further suggests that: 'It may well turn out upon further evidence that intolerance of perceptual ambiguity is related to a broader psychological disturbance of which prejudice...is but another manifestation.' The concept of 'higher tension level', as utilized in this investigation, is assumed to provide that 'broader psychological' basis. From this standpoint, the findings by Lewis & Taylor (1955) and Berg & Collier (1953) that maladjusted subjects tend to have more extreme responses than normal subjects are congruent with the present findings, and can be accounted for on the same theoretical basis.

Whether positive association can be expected between 'intolerance of ambiguity' as defined by extreme response set scores on the P.F.Q. and that same tendency as defined by perceptual measures such as the 'changing figures test' or the 'design preference test' (French, 1955) is a matter for further investigation. Concerning the present study the main objective has been achieved. Extreme response scores, utilized as a measure of intolerance of ambiguity, were found to reveal substantial differences between social groups with different tension levels.

I am indebted to Prof. H. J. Eysenck for his critical reading of this paper in manuscript, and for his valuable comments and suggestions. The material presented in this paper was partially utilized in a Ph.D. thesis submitted by the writer to Cairo University in 1954.

References

Berg, I. A. (1953). The reliability of extreme position response sets in two tests. *J. Psychol.* **36**, 3–9.

Berg, I. A. & Collier, J. S. (1953). Personality and group differences in extreme response sets. *Educ. Psychol. Meas.* **13**, 164–9.

Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educ. Psychol. Meas.* **10**, 3–31.

Eysenck, H. J. (1953). *The Structure of Human Personality.* London: Methuen.

French, E. G. (1955). Interrelation among some measures of rigidity under stress and nonstress conditions. *J. Abnor. Soc. Psychol.* **51**, 114–17.

Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personality*, **18**, 108–43.

Frenkel-Brunswik, E. (1951). Personality theory and perception. *Perception: An Approach to Personality.* Ed. by R. R. Blake and G. V. Ramsey. New York: Ronald.

Lewis, N. A. & Taylor, J. A. (1955). Anxiety and extreme response preferences. *Educ. Psychol. Meas.* **15**, 111–16.

Rundquist, E. A. (1950). Response sets: a note on consistency in taking extreme positions. *Educ. Psychol. Meas.* **10**, 97–9.

Terman, L. M. & Miles, C. C. (1936). *Sex and Personality.* New York: McGraw-Hill.

Winslow, C. N. & Frankel, M. N. (1941). A questionnaire study of the traits that adults consider to be important in the formation of friendship with members of their own sex. *J. Soc. Psychol.* **13**, 37–50.

(*Manuscript received* 1 *July* 1957)

DRIVE STRENGTH AND ADAPTATION TO STRESS

By H. M. B. HURWITZ and J. ROWELL

Psychology Department, Birkbeck College, London

Support for the hypothesis that adaptive responses reduce the drive component in behaviour was found in an experiment involving two groups of rats placed, for different periods, in a water tank. After being detained in a delay chamber, the rats were allowed to escape after making a brightness discrimination. A discrimination reversal test was undertaken in order to establish whether the original learning had been affected by the length of the water-detention period. It was found that a long prediscrimination detention period facilitated reversal learning. This result was taken to support the major contention of the paper, that prolonged exposure to a potential stress situation encouraged the elaboration of drive-minimizing responses and, in consequence, favoured discrimination learning.

I. Introduction

When an animal such as the laboratory rat has been deprived of food, drinking water, a sexual mate, warmth, light, etc., the intensity of its goal-directed behaviour is often said to be related to the extent of the deprivation. Hull (1943) has generalized these observations by claiming that 'drive strength' (D) is a function of the preceding degree of deprivation. As a first attempt at a generalization, this statement has much to recommend it. In the case of food-deprivation at least, the experimental evidence supports the principle (Yamaguchi, 1951; Keller & Schoenfeld, 1950). However, some experimenters (e.g. Kellogg, 1941) have maintained that the repetition of a typical drive-inducing situation leads to modifications in the hypothesized drive-state. In these experiments it appears that continued and repeated exposure to electric shock reduces the motivational efficacy of this stimulus. Consequently, the animal's performance in successive trials of a learning experiment may be conceptualized as the outcome of two factors: increase in habit strength (Hull, 1943) and reduction in 'activation level' (Woodworth & Schlosberg, 1954). However, it appears extraordinarily difficult to obtain clear evidence for the latter part of this hypothesis, i.e. for the suggestion that responses emerge which function to reduce drive level. The present experiment represents a first step towards a solution; we were able to obtain direct evidence that the effect of stress was to encourage the development of a set of responses which ameliorated the stress conditions and minimized typical stress reactions.

Our primary objective of isolating the development of stress adaptive behaviour was undertaken in the following manner: experimentally naïve rats were placed into a water tank for *different* periods and, thereafter, required to perform a brightness discrimination. On making the correct choice, the animals were allowed to escape from the tank by landing on a raised, dry platform. Depending on one's theoretical orientation, two conflicting predictions arise. (*a*) On the basis of the simple equation that duration of stress is correlated with drive strength, a long detention period should intensify the drive to escape from water, raise the reward effect of escape, and increase the 'magnitude of reinforcement' factor for the instrumental responses leading to escape. These factors should combine to result in the formation of a *stronger* discrimination habit than following a shorter water-detention period; a discrimination reversal test should show that the

discrimination requires *more* trials to be reversed after high-drive training than after low-drive training. Further, since in Hull's 1943 theory the magnitude of reinforcement indirectly determines the rate of habit-strength growth, the high-drive condition should lead to fewer acquisition trials. (*b*) The point of view defended in the present paper leads to a different prediction. Increasing the detention period favours the elaboration of adaptive responses which, over a succession of trials, minimize the stress state, i.e. modify drive strength. The result is twofold. First, in terms of the number of acquisition trials required to bring the brightness discrimination habit to its appropriate performance level, there need be no difference between the two groups of subjects. As regards the ease of discrimination reversal we can, as yet, make no definite predictions. Secondly, our position commits us to the prediction that there are *qualitative* differences in behaviour under short and long detention periods, i.e. that the adaptive responses to being placed in water will differ markedly.

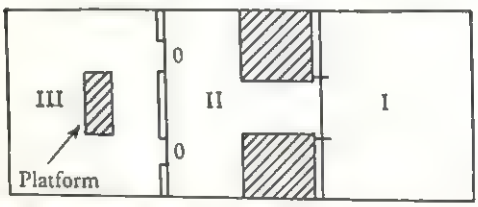


Fig. 1. Plan of the apparatus: I—delay chamber; II—choice chamber; III—escape chamber. Each chamber was divided by a movable partition. Above the partitions separating the choice and escape chamber, two lights were fixed which were switched on in random order during the brightness discrimination trials.

II. The experiment

Subjects. Ten hooded male rats, 3–5 months old at the beginning of the experiment, and drawn from the inbred colony maintained by the Department of Psychology, were used in the present study. The animals were allotted randomly to each of the two experimental groups.

Apparatus. A 3 ft. long water tank, 14 in. deep and 14½ in. wide, was divided into three sections by two partitions (see Fig. 1). The first partition was positioned 15 in. from one end of the tank. It contained a single shutter, 8 × 6 in., placed centrally, at water level. The animal could swim from the first chamber (delay chamber) into the second chamber (choice point) when the shutter was opened by the experimenter. The second partition was placed 12 in. from the other end of the tank, and contained two shutters, each 4 × 6 in., at the right and left extremes. An enclosed lamp-holder was fastened above each shutter. In this way, the doors could be illuminated during brightness discrimination trials. When the shutters were opened, the animal could swim from the choice point into the escape compartment. The escape chamber was 9 in. long and contained a floating platform onto which the animal scrambled to escape from the water. The water was 6½ in. deep and was kept at 65–69° F.

Procedure. Both experimental groups received the same preliminary training. (1) Each animal was put, in turn, into the water in the escape chamber. As soon as it had scrambled onto the escape platform it was removed from the apparatus and returned to its home cage. This routine was repeated three times. (2) On the next three trials, the animal was placed in the *choice point* chamber, both shutters being open and a free choice of escape routes thus being available. (3) For the next three trials all the shutters were kept open; the animal was placed in the *delay chamber* and could effect its escape by swimming through all three chambers.

After preliminary training, the experimental conditions were as follows. Each group had to swim through the three chambers and select the brighter of the two shutters in order to effect an *immediate* escape. A wrong choice merely delayed escape by a few seconds, since the animal could swim to the other side of the partition and clamber through the correct, open shutter. Group A was detained in the delay chamber for 10 sec. during training trials; group B was detained for 40 sec. Twenty-five training trials were given: approximately five trials spaced by 10 sec. per session.

At the conclusion of the training period, a further twenty-five trials were run. On these occasions the correct shutter was the darker of the two (i.e. habit reversal) and each group was detained for 25 sec. in the delay chamber. Records were obtained of (*a*) the total number of errors made by each animal during training and reversal trials, (*b*) the time taken from the opening of the first shutter to the animal's arrival in the escape chamber, and (*c*) the manner of dealing with the detention period.

III. Results

Error and time scores during training trials

In terms of the total number of *errors* made by both groups during training trials, no difference in the rate of discrimination learning could be established. A test for difference in learning rate as measured by reaction times also failed to indicate a statistically significant difference between group A and group B.

Reversal training

A U-test (Wilcoxon, 1945) analysis of the error scores during reversal training indicated a difference between group A and group B which was statistically significant at the 0·02 level. Group A (short-detention period during training) required more trials to reverse the discrimination than group B. The median score for group A was thirteen errors and for group B, seventeen errors.

Behaviour during detention—adaptation responses

During *preliminary* training trials, all animals exhibited typically stressful behaviour, i.e. great activity, urination, rapid heart beat, whining, defecation, etc. Mowrer & Viek (1948) suggest that this kind of behaviour is highly analogous to 'fear from a sense of helplessness'. On being placed in the water, the animals jumped up walls, splashed and thrashed about frantically and attempted to hold fast to any ledge. In the *training* trials proper, these severe stress signs were greatly reduced and, at the end of the training period, had almost completely disappeared. Further, towards the middle of the training period, it became clearly noticeable that group A was very much more *active* in the delay chamber. This activity consisted of vigourous swimming and clawing at the exit door. Group B, on the other hand, developed a tendency to remain perfectly motionless and apparently relaxed for as long as 30 sec. During the final few seconds, group B became active, swimming gently towards the opening door. All five animals in this group showed this behaviour.

During *reversal* training, when both groups were detained in the delay chamber for 25 sec. before being admitted to the choice chamber, *group A developed the adaptive behaviour previously observed in group B, whereas group B reverted to the pattern exhibited by group A throughout the final portion of the training period.*

IV. Discussion

Turning first to the prediction derived from Hullian theory, we see that our results are at variance in a number of ways. Thus, the longer detention period did *not* influence the rate of learning the original discrimination, i.e. a long detention in water does not necessarily raise drive-level, or increase the magnitude of reinforcement for escape responses. And again, the reversal test failed to show that longer detention in the original

learning trials resulted in a stronger habit, i.e. that it would require more trials to acquire the reverse discrimination following long detention than following short detention in the water. But the case made out in the 'Introduction' for assuming that when animals are subjected to a stressful situation they tend to develop adaptive responses which serve to minimize the drive component in behaviour, seems well supported by the present evidence. In the present experiment we have been able to pin-point the adaptive responses which displace or interfere with the original stress reactions and have, thereby, added substance to Kellogg's (1941) suggestion that the motivational efficacy of stress-arousing stimuli becomes attenuated. There are two things worth noting: first, the adaptive responses develop rather slowly, i.e. they reached their effectiveness only after a number of trials. Secondly, the adaptive responses are, in their turn, modifiable and adapted to the nature of the task. Thus, as already indicated, when the detention period for group B was changed from 40 to 25 sec., the adaptive responses changed, i.e. instead of remaining motionless for about 30 sec., the animals engaged in calm unruffled paddling. Our reason for calling these responses 'drive reducing' or 'drive modifying' derives from the more generally accepted description of the role of drives in regulating behaviour. It is said that behaviour varies both in intensity and directness, and that drive relates to the former quality, whereas 'motive' includes both qualities (see Marx, 1956; Hebb, 1955). The function of the responses here singled out for examination clearly affect the *intensity* of the performance, and in fact reduce it. However, we have also shown that, in addition to modifying the intensity of behaviour, the adaptive responses influence the quality of learning. Thus, group B's superior performance on the reversal test seems intimately related to their wider adaptive response repertoire. This inference is supported by the observation that during the initial reversal trials, group A elaborated *new* adaptive patterns prior to exhibiting any signs of learning the reversed discrimination.

The experiment reported in this paper is part of a B.Sc. (Special) thesis submitted by Miss Rowell in 1955. The construction of apparatus was financed by the Central Research Fund, University of London, and the Department of Psychology, Birkbeck College. We would like to express our appreciation to Mr S. M. Evans, M.B.E., for supplying a photographic record of the experiment, to Miss D. R. Bolas for help in the preparation of the manuscript, and to Mr H. Tajfel for invaluable criticisms.

References

Hebb, D. O. (1955). Drives and the C.N.S. (Conceptual Nervous System). *Psychol. Rev.* **62**, 243–54.

Hull, C. L. (1943). *Principles of Behavior.* New York: Appleton Century.

Hull, C. L. (1951). *Essentials of Behavior.* New Haven: Yale University Press.

Keller, F. S. & Schoenfeld, W. N. (1950). *Principles of Psychology.* New York: Appleton Century.

Kellogg, W. N. (1941). Electric shock as a motivating stimulus in conditioning experiments. *J. Gen. Psychol.* **25**, 85–96.

Marx, M. H. (1956). Some relations between frustration and drive. In *Nebraska Symposium on Motivation.* University of Nebraska Press.

Mowrer, O. H. & Viek, P. (1948). An experimental analogue of fear from a sense of helplessness. *J. Abnorm. Soc. Psychol.* **43**, 193–200.

Woodworth, R. S. & Schlosberg, H. (1954). *Experimental Psychology.* London: Methuen.

Yamaguchi, H. (1951). Drive as a function of hours of hunger. *J. Exp. Psychol.* **42**, 108–17.

(*Manuscript received* 8 *March* 1957)

THE MEANING OF 'INTELLIGENCE' TO LEBANESE TEACHERS

By JACK D. KEEHN and E. TERRY PROTHRO
Psychology Department, American University of Beirut

The question of the meaning of *intelligent* for English-speaking Lebanese teachers was explored by having twelve teachers rate students in nine classes on eleven scales, including a scale of intelligence. A factor analysis was made of these ratings. It was found that judgements of intelligence covaried with judgements of conscientiousness, thoughtfulness, persistence, and emotional stability, and that such judgements were relatively independent of judgements of cyclothymia and friendliness. The results of this analysis were similar to the results of comparable American studies. This investigation would seem to point up the value of the factor analytic approach to the problem of meaning.

I. Introduction

In an investigation of the predictive efficiency of certain 'culture-free' intelligence tests including Cattell's *Culture-Free Test* (1944), Raven's *Progressive Matrices* (1938) and the French D48 (1952) version of the *Dominoes Test* (cf. Vernon & Perry, 1949), Keehn & Prothro (1955) used as one of their criteria the judgements of bilingual Lebanese teachers. There was little relationship between scores on the tests and intelligence ratings by the teachers. This lack of relationship could be interpreted as reflecting adversely on the efficacy of the tests, but it might also raise some question of the manner in which the teachers use the term *intelligent*. Teachers' ratings with respect to intelligence cannot be used as criteria in cross-cultural investigtions unless teachers in different cultures (or sub-cultures) are using the term similarly. If, for example, teachers in one culture judged an energetic, loquacious person to be intelligent while those of another culture thought of such a person as rather unintelligent, then ratings could hardly serve as criteria in both cultures.

Empirical data on the manner in which *intelligent* is used can be gathered by having judges rate a number of individuals on several different traits. Factor analysis of the ratings will then show the patterns of interrelations between trait names. If it were found that judgements of *cheerful* and *dominant* covaried regularly with those of *intelligent* for different raters and ratees in one culture, then some inference could be made about the 'meaning' of the term *intelligent* in that culture.

Osgood (1952, 1954, 1955) has made extensive use of the factor analytic approach to the problem of meaning. He maintains (1952, p. 227) that 'judgment can be conceived as the allocation of a concept to an experimental continuum' and that 'many different experiential continua...are essentially equivalent'. Thus judging a pupil's intelligence involves treating him as the 'concept' and treating *intelligence* as an 'experiential continuum'. Moreover, other continua, such as *thoughtfulness*, might be shown by analysis to be essentially equivalent to the *intelligence* continuum. In this fashion, then, the meaning of *intelligence* can be determined by locating the term in a 'semantic space' which is described in terms of the dimensions emerging from factor analyses.

Data are already available on the factor analysis of rating scales with American raters and ratees. Cattell's work on the description of personality, although directed toward the

discovery of source traits, revealed a great deal about the use of *intelligent* by his raters when he demonstrated that one of his source traits included *intelligent*, *conscientious*, *thoughtful*, and *emotionally stable* (1946, p. 313). That is, American raters judging many different individuals showed a definite tendency to rate subjects as *conscientious* and *thoughtful* whenever they rated them *intelligent*. The 'meaning' of *intelligent*, then, has much in common with the 'meaning' of these other terms, at least for these American judges.

In order to determine something about the 'meaning' of *intelligent* to English-speaking Lebanese teachers, twelve such teachers were asked to rate their pupils on a series of scales. The scales were then factor analysed to see what general tendencies there were to respond similarly to different trait names.

II. Method

The study was carried out at the International College located in Ras Beirut, Lebanon. The pupils of this institution are Arabs, but English is the language of instruction for most of the courses. The instructional staff is largely made up of bilingual Arabs, and the administration is American. The word 'College' in the title of the institution is used in the European sense. That is, work is offered up to the American university sophomore level.

Teachers in this college were asked to rate pupils on the following traits: friendliness, intelligence, emotional stability, dominance, cyclothymia, sympathy for others, conscientiousness, co-operativeness, persistence, conventionality, thoughtfulness, and trustfulness. The list was based in part on Cattell's factorial studies of personality ratings (1946, pp. 313 ff.). For each trait or characteristic, a five-point scale was constructed with statements indicating the five steps. The scale for intelligence, for example, read as follows: 'highly intelligent'; 'above average intelligence'; 'average intelligence'; 'rather below average intelligence'; and 'unintelligent'. The scale for cyclothymia placed 'very cheerful and talkative' at one extreme and 'very quiet and seclusive' at the other extreme. The word *cyclothymia* did not appear on the scale.

Ratings were made by twelve teachers on nine classes of students. These twelve teachers were among those participating in the previous study of intelligence tests (Keehn & Prothro, 1955). Teachers rated from twenty-eight to thirty-one pupils in each class.

III. Results

Most of the teachers omitted the item on trustfulness, so this item was excluded from the analysis. The other eleven ratings on each class were intercorrelated separately for each teacher. The fourteen resulting correlation matrices were factor analysed separately. In each case virtually no significant residuals remained after the extraction of two centroid factors. These factors were rotated independently and blindly for each analysis. Rotations were done graphically by passing one axis through the most predominant cluster and retaining the second orthogonal to it. The loadings so obtained are shown in Table 1. Inspection of Table 1 reveals that there is considerable similarity in the factorial patterns found in the several matrices. The summary presented in Table 2 reveals the patterns quite clearly. Factor 1 seems to be defined in terms of *thoughtful*, *conscientious*, *intelligent*, *persistent*, and *emotionally stable*. Factor 2 is defined by *cyclothymic* and *friendly*. The terms *dominant*, *co-operative*, and *sympathetic* have substantial loadings on both factors.

Table 1. *Rotated factor loadings from 14 correlation matrices*

Teacher	A		B		B		C		D		E		F	
Class	10*a*		8*b*		8*c*		7*b*		8*b*		7*b*		4*d*	
Trait	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2
Friendly	03	93	30	65	14	76	24	88	67	12	80	35	03	83
Intelligent	65	31	33	48	66	46	58	46	83	26	53	63	77	43
Emotionally stable	42	−13	74	−11	72	−31	44	07	88	−12	52	01	63	13
Dominant	36	55	35	75	−02	72	46	−08	12	68	−03	85	66	40
Cyclothymic	12	87	−05	61	−20	77	02	87	−01	73	10	78	−02	60
Sympathetic	08	75	53	24	40	72	57	41	71	−09	80	−35	02	80
Conscientious	75	−08	93	13	89	−06	80	34	91	15	59	−46	61	22
Co-operative	20	61	66	40	47	68	67	21	89	−14	83	−05	06	78
Persistent	74	41	65	44	83	33	74	00	86	25	42	58	65	31
Conventional	35	18	00	82	09	70	61	18	−04	−67	06	63	54	09
Thoughtful	90	−05	83	28	86	−01	72	−08	83	20	87	02	81	37

Teacher	G		H		I		J		K		K		L	
Class	4*d*		9*a*		9*a*		8*a*		10*a*		11*a*		10*b*	
Trait	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2	F_1	F_2
Friendly	08	67	52	22	45	77	17	73	14	61	00	94	14	78
Intelligent	86	15	83	16	88	07	76	29	74	23	86	07	88	21
Emotionally stable	73	−34	70	−06	72	22	64	−09	40	13	40	02	49	−06
Dominant	−13	85	04	68	63	11	−11	38	89	04	74	07	37	51
Clyclothymic	−15	77	−03	54	22	82	−06	68	16	77	−05	79	08	69
Sympathetic	55	−09	63	−03	58	39	61	55	−32	66	03	71	23	64
Conscientious	75	−18	78	45	87	−02	84	−11	67	−03	71	10	74	−04
Co-operative	51	29	90	−04	81	21	56	49	24	66	42	70	−02	67
Persistent	25	58	79	15	91	−03	40	07	81	15	67	−15	44	11
Conventional	−37	51	01	71	−08	42	52	11	72	−07	91	−05	61	31
Thoughtful	79	−38	72	40	88	−08	79	12	84	25	88	−02	80	−05

Table 2. *Summary of data on loadings from 14 correlation matrices*

Trait	No. loadings on $F_1 > 0{\cdot}60$	No. loadings on $F_2 > 0{\cdot}60$	Median F_1 loading	Median F_2 loading
Friendly	2	11	15	75
Intelligent	11	1	77	27
Emotionally stable	8	0	63	−06
Dominant	4	6	35	53
Cyclothymic	0	13	01	77
Sympathetic	4	6	54	48
Conscientious	13	0	77	−03
Co-operative	6	6	53	45
Persistent	10	0	71	20
Conventional	4	4	22	25
Thoughtful	14	0	83	01

IV. Discussion

From our results it would appear that the circumstances under which our teachers would use the term *intelligent* have a great deal in common with those under which they use *thoughtful*, *conscientious*, *persistent*, and *emotionally stable*. The fact that ratings on these five trait names tend to vary together gives us some insight into the situations which control the use of *intelligent*. It also appears that judgements of *cyclothymic* and *friendly* have virtually no relationship to judgements of *intelligent*. The variables controlling our teachers' use of such terms as *friendly*, then, are not those controlling their use of *intelligent*.

It is interesting to note that the two factors which we have extracted in this analysis are similar to two of Cattell's factors. The factor which he refers to as 'general mental capacity' includes 'intelligent', 'conscientious', 'thoughtful', 'stable emotionally', 'intellectual', and 'wise'. Apparently this is very close to our Factor 1. The factor which

he refers to as 'cyclothymia' includes 'outgoing', 'good-natured', 'friendly', 'cheerful', and 'trustful'. This factor is quite similar to our Factor 2.

The similarities between the results of our analysis and that of Cattell would suggest that in Lebanon as in America the variables controlling the use of *intelligent* are frequently the same as those controlling the use of *conscientious*, *thoughtful*, and *emotionally stable*. There is, then, considerable overlap in the 'meaning' of these terms in both countries. No evidence was found for a unique use of *intelligent* on the part of bilingual teachers. Consequently, the low correlations between intelligence ratings assigned by Lebanese teachers and performance on 'culture-free' intelligence tests, Keehn & Prothro (1955) can be interpreted as reflecting unfavourably on the validity of those tests.

This investigation points up an important application of the factor analytic approach to 'meaning'. Osgood's suggestion that this technique is useful for studying cross-cultural problems in communication is borne out. Such analysis provides an approach to the perennial problem of whether 'meanings' are the same in investigations which involve the verbal behaviour of persons in different cultures.

References

Cattell, R. B. (1946). *Description and Measurement of Personality.* Yonkers-on-Hudson: World Book Co.

Keehn, J. D. & Prothro, E. T. (1955). Non-verbal tests as predictors of academic success in Lebanon. *Educ. psychol. Measmt.* **15**, 495–8.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychol. Bull.* **49**, 197–237.

Osgood, C. E. & Luria, Z. (1954). A blind analysis of a case of multiple personality using the semantic differential. *J. Abnorm. Soc. Psychol.* **49**, 579–91.

Osgood, C. E. & Suci, G. J. (1955). Factor analysis of meaning. *J. exp. Psychol.* **50**, 325–38.

Vernon, P. E. & Parry, J. B. (1949). *Personnel Selection in the British Forces.* London: University of London Press.

(*Manuscript received* 5 *September* 1957)

PUBLICATIONS RECENTLY RECEIVED

Body and Mind in Western Thought. By JOAN WYNN REEVES. Harmondsworth: Penguin. 1958. Pp. 403. 5*s*.

Psychologists, as Dr Reeves points out in her introduction, fall into two groups, those who 'put aside all questions about body and mind as philosophically senseless and scientifically a waste of time', and those who hold that they may provide a key 'to better understanding of psychological theory and practice'. Her book will certainly be welcomed by the latter group, and may well interest and even convert some of the former.

As Pelicans go this is a big book. It is, too, a finished book, admirable in its style, and obviously compiled with much loving care. It consists of two parts. The first is a long historical introduction of nearly two hundred pages by Dr Reeves herself; the second, a selection of extracts from Western thinkers ranging from Hippocrates and Plato to Wundt and William James. A set of eighteen plates of illustrative works of art (ironically enough sandwiched in among extracts from Plato who would have disapproved) contrast with the abstract arguments of the thinkers and remind the reader of the concrete variety of human nature and the vicissitudes of human culture. This indeed is Dr Reeve's striking merit and achievement. She has a keen sense of life and complexity and an awareness of the tides of historical change. She sees every stage of long debate on mind and body set against a background of conflicting cultural and social forces. She brilliantly conveys the emotional overtones which have made the whole topic much more than a mere matter of evidence and logic. The problems of body and mind 'grew so entangled', she points out 'partly because of the age-long identifications of matter with the sensory (and sensual), the emotional (and licentious), the enclosing (the soul's or devil's container), the inert (unresponsive), or the spatially relatable (soon to mean "the determined").... "Mind" by contrast was apt to connote the thinking (and pure), the rational (and controlling), the incarcerated (but immortal), the vital (and spontaneous), the unextended (therefore free)' (p. 74). These identifications coloured the thinking of the seventeenth, eighteenth and nineteenth centuries, and indeed are partly responsible for 'the chaos from which we have by no means emerged'. Again and again just because of the breadth and richness of her approach Dr Reeves provides the reader with illuminating insights which a 'bare bones' account of mind–body theories would have missed.

She is perhaps for that very reason less successful in delineating the main lines of the development of theories about minds and bodies—the gradual purification of animism from its first superstitious beginnings to its spiritual culmination in scholastic philosophy (though Dr Reeves documents the obverse demonology and quotes from the *Malleus Maleficarum*); the majestic and inexorable progress of the mechanical theory of the physical universe through its nineteenth-century zenith to its twentieth-century dénouement; the development of the biological sciences and their gradual expulsion of vitalism; the rise from Descartes onwards of the concept of consciousness, hardly known to the ancients but becoming in the early modern period the key concept of psychology and a crucial issue in philosophy; the overthrow of this concept from its central position as a result of new forces; and finally the contemporary confrontation of a physics quite undogmatic about the nature of 'matter' and a psychology equally undogmatic about the nature of 'mind'. There is in the whole story a drift and a trend; there is, I would think, a grand movement which transcends local circumstances. All this I find somewhat blurred in Dr Reeves's 'projection' by the very intricacy and subtlety of her detail. For that reason her book is more likely to be appreciated by the sophisticated reader rather than by the beginning student in psychology, for whom a clear-cut, even if over-simplified, framework is desirable.

About her own introduction Dr Reeves is engagingly modest. It is 'no more than a sketch map drawn to the editor's projection, to be thrown away, or at most employed as a lamp-shade'. Dr Hughlings Jackson when he bought a book which was too fat for his pockets used to tear off the covers and then split the book in two, much to the surprise of the bookseller in whose presence this act was performed. Few of Dr Reeves's readers are likely to do likewise with her Pelican; and if they did they would not throw her half away. No doubt there will be some who wish to modify her 'projection', for projections are subjective things. My 'projection' would give a larger place to eastern influences in western thought. From the time Alexander conquered the east, oriental religions and philosophies bulked larger and larger in the west, and the new civilization that arose from the ruins of the Roman Empire was an amalgam of

east and west. The west without Christianity and its eastern component would not be the west as we know it, nor would western thinkers have psychologized as they did. Neither Pythagoras at one end nor Jung at the other end of the long succession can be understood without the influence of eastern thought, and in so far as the orient became domesticated in the west it became an integral part of the western mind.

In modern times Dr Reeves omits both Marx and psychic research from her map. They both seem to me to be sufficiently important to include. Marx had perhaps as much as anyone to do with the rise of a biosocial view of mind and personality though his influence is not always acknowledged. In any case his contemporary dominion is so widespread that he ought not to be ignored. As for psychic research, it cannot be denied that it has coloured the views of at least some psychologists of the front rank. If its evidence is acceptable, the problems it raises are formidable.

Dr Reeves's anthology has been skilfully done and is just the right mixture of solid fare and idiosyncratic condiment to be really palatable. Obviously two hundred pages cannot contain everything, and it would be pointless to complain of omissions. All goes reasonably well until the nineteenth century. Then the Pelican loses its appetite and shows signs of satiety. Thirty-six pages from Braid, J. S. Mill, Hering, Wundt and William James is all it can manage and this is not doing justice to the fare provided by the nineteenth century. I can see no solution to this but to ask Dr Reeves to write another book on 'Body and Mind in nineteenth and twentieth century Western Thought'. Meanwhile she must be thanked for a book which will add a new dimension to the thinking of many present-day students of psychology. Perhaps the simultaneous publication of Heisenberg's *The Physicist's Conception of Nature* which does for physics just what Dr Reeves has done for psychology (combining a commentary with an historical anthology) will convince even tough-minded scientists on the necessity of her enterprise.

L. S. HEARNSHAW

Educational Psychology and Children. By K. Lovell. London: University of London Press. 1958. Pp. 272. 18*s*.

A moderately priced text-book cannot afford to be lengthy. Dr Lovell's book succeeds in covering a wide field for its moderate length. The first nine chapters range over general psychology, including motivation, abilities, personality, perception, maturation and learning. The second nine chapters have particular educational relevance and cover educational guidance, backwardness and maladjustment, as well as a sketch of child development on the Gesell pattern. (Two concluding chapters, on physiology and statistics, are rather awkward appendages, being too condensed to do more than bewilder and dismay a student.) The author has performed a considerable task of abstraction to cover this range of topics without omitting contrasting points of view: thus, his nine-page chapter on 'Motivation' describes instinct theory, psychoanalysis, social theory, field theory and learning theory. Inevitably the text is indigestible at times. Only a much longer—and consequently more expensive—book could have provided this width of coverage with full detail and illustration.

There is plenty of detail. The book indeed is written rather in the style of the journal article, with footnotes identifying each contribution with author's name and reference. It is written on the basis of published research, with over 400 references in the text. Two-thirds of these references are to books or papers published in Britain (or jointly in Britain and U.S.); nearly half the references are dated 1950 or later; and over fifty refer to papers in post-war numbers of the *British Journal of Educational Psychology*. This emphasis on research will be welcomed by those who wish to depart from the tradition of writing a psychological text as literature. There are no anecdotes here, nor any illustrative John or Mary whose artificial case-history presses home a point. The presentation is clear enough, but dull: by Flesch's formulae, its Readability Index is 44 but its Human Interest score is only 8 (on a 100-point scale). Whether a text-book should be full of human interest is a debatable point, for the chatty introduction and the inspiring conclusion (to be found, for example, in Hughes and Hughes's *Learning and Teaching*) do not have the same appeal to the student working for his examination as to the general reader. However, an illustrative example may earn the valuable space it occupies by adding to the clarity and memorability of a difficult point; and though some examples are given here, in general the treatment is highly abstract.

Consequently, I doubt if this book will convert many teachers to a recognition of the value of psychology in education. It will give students for whom it is prescribed a clear idea of recent psychological research; and if they are not impressed by the relevance of what they read to the practical task of classroom teaching, it may not be so much the fault of the author as a reflexion of the bias of recent educational research in this country. Where the relevance of a chapter to education seems slight, the author has

included a section on practical points to note, but these do not always arise directly from the research which he has been describing. The author does not follow the common practice of other text-books in giving at the end of each chapter a list of practical exercises for the student, or of points to discuss, or problems in which to apply the principles discussed. Had he done so, he might more readily have escaped the accusation of writing a book designed to prepare students for a written examination, rather than to prepare them for their practical work of teaching. Nevertheless, among those who will find this book of value, not the least will be teachers of psychology, for whom it will provide a brief, competent and manageable review of recent work in both educational and general psychology. JOHN NISBET

Psychotherapy of Chronic Schizophrenic Patients. Ed. by CARL WHITAKER. London: Churchill. 1958. Pp. ix+219. 35*s*.

A New Approach to Schizophrenia. By J. I. STEINFIELD. London: Hutchinson. 1958. Pp. 159. 21*s*.

Psychotherapy of Chronic Schizophrenic Patients takes the form of an informal discussion among a small group of American psychiatrists who meet annually to exchange views on the diagnosis and treatment of schizophrenics. The book consists of eight chapters each comprising a discussion of some aspects of schizophrenia and led in turn by one of the eight members of the group. The current American preoccupation with communication in psychiatry, represented by the group's belief 'that people communicate more to one another than what they say', results in every word spoken in discussion being faithfully reproduced in print and, at times, this proves tiresome to the reader. As the group are all experienced psychotherapists, a large measure of agreement is reached on the most suitable psychotherapeutic approach towards schizophrenic patients. This concensus is, however, noticeably absent in the group's discussion of the diagnosis of the schizophrenic condition. Indeed it is admitted in the summary following the chapter on 'Diagnosis and Prognosis' that the foregoing discussion is 'somewhat confused and disorganised'. When, in succeeding chapters, the group pool their therapeutic experiences one is very impressed by their ability to understand, communicate with and apparently relieve schizophrenic patients. The general aetiological viewpoint expressed is that the schizophrenic patient is a resultant of a basic family disorder, and two of the chapters centre specifically on the treatment of the patient within the family group.

The only newcomer to the panel, invited specially for this conference, is the anthropologist Gregory Bateson, and it is his contribution on 'Schizophrenic Distortions of Communication' which provides the most stimulating discussion and goes nearest to producing a coherent theory of the schizophrenic process. Bateson presents an interesting account of his 'Double Bind' theory where the basic disorder of schizophrenia is seen to be a distortion in the classification of messages. His thesis is well presented and adds support to the views of Schilder, Cameron, Arieti and others who have also been interested in schizophrenia as a disorder of communication. Bateson also raises such worthwhile questions as whether one can validly speak of 'unconscious judgement' and whether we have any right to assume purposive behaviour behind the schizophrenic's process. Unfortunately, such questions are largely ignored due to the group's preoccupation with psychotherapy which fundamentally rests on the assumption that schizophrenic behaviour is purposeful to the patient. The book is undoubtedly well worth reading by anyone interested in the fascinating problems raised by the schizophrenic patient, particularly those with a therapeutic orientation.

While Whitaker and his group consider that the schizophrenic patient has been starved of affection within the family group, Steinfeld in *A New Approach to Schizophrenia* suggests that the schizophrenic has been literally starved. In his opinion, all schizophrenics suffer a severe early hunger trauma during the first few weeks of life at a time when the infant has a purely vegetative existence. As a result of this trauma the individual is conditioned to abnormal vegetative reactions which inhibit the normal supremacy of cortical functions in later development. The author's rather startling approach to treating the schizophrenic is simply to reproduce the original hunger trauma directly by starving the patient, or indirectly to recapitulate the original state of infantile acidosis by having the patient inhale acetone while under an anaesthetic produced by E.C.T. or Pentathol. Steinfeld reports remarkable improvements in patients who have been treated in such a way, but is careful to add that, in his view, there is no cure for schizophrenia, the therapeutic aim being limited to increasing the periods of remission. He suggests that we must turn our thoughts to the prophylactic measures which can be taken to ensure that the predisposing hunger trauma does not occur. In this direction, he recommends a nation-wide campaign to encourage

unlimited feeding on demand to ensure adequate early nourishment. The book contains thirty-nine illustrative case histories of schizophrenic patients who all demonstrate the typical vegetative history which our author believes may be found only in schizophrenics. In almost half the total group of 134 patients stated by the author, early histories were unobtainable, but it is nevertheless assumed that the missing cases 'would probably have shown the specific early history had we been in a position to study it'.

The author closes with a list of the elements which constitute a typical vegetative history and perhaps the chief merit of the book is that the hypothesis presented is accessible to objective validation. One feels doubtful, however, considering the apparently world-wide incidence of schizophrenia, that such a specific and simple explanation of this condition is likely to be valid.

ANDREW MCGHIE

Attempted Suicide: its Social Significance and Effects. By E. STENGEL and NANCY G. COOK. Maudsley Monographs No. 4. London: Chapman and Hall. 1958. Pp. 136. 25*s*.

The range of suicidal behaviour extends from suicidal threats and gestures at one pole to the completed act at the other. Previous studies of attempted suicide have been made chiefly with the intention of contributing to the knowledge of suicide. Professor Stengel and his colleagues have broken new ground by studying the problems peculiar to attempted suicide itself, in addition to those it has in common with suicide. In their definition of a suicidal attempt they included slight injuries with no danger to life, provided that these appeared to be consciously aimed at self-destruction: and in an investigation of more than 600 cases admitted to London hospitals develop their thesis that the suicidal attempt should be viewed as a social behaviour pattern in its own right. A parallel study of 117 cases of suicide indicated that the social element in suicide, though not absent, is much less manifest.

Much evidence is adduced to justify the focusing of attention on the life-preserving tendencies inherent in suicidal attempts which are considered by the authors to be no less significant than the aggressive-destructive trends. They point out that a comparatively small proportion of those who have attempted suicide finally kill themselves, and that only a minority of suicides have attempted suicide before. They claim that most suicidal attempts seem to take place within the individual's social field and in such a way that the intervention of others seems possible or even probable. The outcome depends therefore not only on the strength of the individual destructive drives but also on the response of the human environment to this social crisis.

Gestures and threats of suicide are well-recognized and common ways of manipulating the environment to the individual's own advantage. In most if not all suicidal attempts Stengel and his colleagues identify a similar social orientation, an 'appeal element' or call for help. The suicidal individual is usually unaware of this appeal element which makes it all the more potent in eliciting the guilt feelings which so often characterize society's response.

Stengel and his colleagues also describe an 'ordeal character' of the suicidal attempt, a testing or proving of the balance of libidinal and destructive impulses in the self the outcome of which the subject appears, usually fatalistically, to accept.

The variety of social effects of suicidal attempts, to which few previous writers have given attention, are described and emphasized as being as significant as the social factors whose etiological importance is widely accepted. The resulting modifications of human relationships, the reactions of mourning and reparation in relatives, and the punitive social reactions expressed in the present law of England on attempted suicide, are illustrated and discussed. Few would disagree that in assessing the suicidal risk and in organizing appropriate treatment the reactions of the social group must be taken into consideration, and that discovery of particular social factors which are important in causation, for example, isolation, has important implications for the prevention of suicidal acts.

It is probable that the appeal element in suicidal acts is not so constantly present as Stengel and his colleagues claim: it is, for example, at the least much less conspicuous in the suicidal attempts of old people than in those of young hysterics. The authors may also go too far in their suggestion that those who have attempted suicide and those who have committed it constitute two different though overlapping populations: they themselves point out that the incidence of suicide in those who have attempted suicide is high compared with the incidence of suicide among the general population. These doubts and the new vistas opened by the authors in their elucidation of the social setting should stimulate further research. The major hypotheses are clearly stated and may be tested. Throughout, the monograph is lucidly and persuasively written, and alike in its careful factual reporting and in its original theoretical formulations is an important contribution both to psychiatry and to sociology.

I. R. C. BATCHELOR

The Neuroses and Their Treatment. Ed. by E. PODOLSKY. New York: Philosophical Library. No date given. Pp. xiv+555. $10.00.

This collection of thirty-seven papers is intended to provide a comprehensive survey of contemporary (American) theory and practice concerning the diagnosis and treatment of neurotic conditions. Consistent with this aim, the patients described range from young infants to the aged and their disorders from relatively mild emotional reactions to everyday stresses to severe and complex neurotic and psychosomatic conditions. The methods of therapy described include varieties of psychotherapy, psycho-pharmacological procedures, physical treatments and psycho-surgery.

The intended reader is less well defined. Psychiatrists and clinical psychologists are likely to consider the discussions insufficiently detailed and will certainly regret the complete lack of bibliographical references and an index. Companion volumes in this series, on literary, philosophical, artistic and general scientific topics appear to be aimed at the 'intelligent layman', but he would experience considerable difficulty with the technical vocabulary of the present volume. Possibly the work is best suited to the needs of physicians in general practice for whom several of the papers appear to have been specifically written.

The papers constituting the chapters were culled from publications in psychiatric journals by 'eminent authorities in the field on all phases of the neurosis'. Their very eminence dictates the nature of the treatment which is mainly that of the 'authoritative clinical report' with its attendant weaknesses. In most chapters a theoretical viewpoint or interpretation is discursively expressed and illustrated by reference to selected clinical histories described as typical. No controlled experimentation is reported even when the title of a chapter refers to experimental studies. Similarly no statistics of greater sophistication than a percentage are presented except in one chapter by Miles, Barrabee and Finesinger on the evaluation of psychotherapy. Following a review of the literature, these authors make a thorough attempt to assess the effects of psychotherapy on sixty-two cases of anxiety neurosis treated in a general hospital.

Many chapters are devoted to discussions of aetiology and most of these assume the general validity of psycho-analytic theory though, frequently, minor or major specific modifications or additional hypotheses are proposed. Similarly, almost all discussions of psychotherapy assume the efficacy of the psycho-analytic techniques of free association and manipulation of the 'transference' situation, though theoretical explanations of these effects vary. This is even true of one chapter by Zuger who proposes, on the basis of a somewhat metaphysical theory, that the aim of therapy is to make the conscious unconscious.

Several of the chapters devoted to therapy are concerned with the use of 'tranquillizer' drugs. These tend to display an over-optimistic attitude and lack of objective evaluation of therapeutic value. Bayart, in an extreme example, states: 'The person treated with hydroxyzine retains full consciousness of the stimuli, affective or otherwise, that released this disturbance. He reports their nature and their intensity but his reactions have become those of a completely balanced individual.' Concerned with psycho-pharmacological treatment of a different type is Meduna's description of his development of carbon dioxide therapy, a fine example of the 'tough-minded' approach to clinical research.

The main interest in this volume, at least for the reviewer, lies in the many excellent descriptions of case-histories illustrating the wide range of symptomatology and behaviour which may be described as neurotic. As always, there is also much to interest a psychologist in the speculations of experienced clinicians and when these are based on the careful observations of, for example, a Spitz, he is stimulated to reassess the aptness of his own theoretical frame of reference to the data described.

H. GWYNNE JONES

The Testing of Negro Intelligence. By AUDREY M. SHUEY. Lynchburg, Virginia: J. P. Bell. 1958. Pp. xv+351. No price given.

As the author herself says, this book might properly be called *A Comparative Study of the Intelligence of American Negroes and Whites.* Her address, *Lynch*burg, Virginia, and her acknowledged indebtedness to her students in the '*Racial* Psychology Classes' (at Randolph-Macon College) are sufficient indicators of her motivation and purpose in writing a book on this subject of lively debate in the United States.

The book is a painstaking and careful review of approximately 240 experimental studies, over 44 years, of negro intelligence, of which all but two were carried out in the States. The studies chosen deal primarily with intelligence as measured by *standard* psychometric tests, some sixty of which are involved. Dr Shuey

groups these studies into eight primary divisions to each of which she devotes a chapter. After an introductory chapter, she continues with one on Young Children, that is, children primarily in nursery and kindergarten schools. There follows chapters on School Children, High School and College Students, the Armed Forces, Deviates, Delinquents and Criminals, and Racial Hybrids; next comes a chapter on Selective Migration; then a final chapter summarizing the findings and giving the author's conclusions.

The results of the various studies in the several chapters are conveniently presented in tables giving the author, date and place of each study, the number of subjects involved, their ages, the method by which they were selected and the results. The latter are clearly distinguished from the comments and conclusions of the investigators thus enabling the reader to assess the data without prejudice. When white subjects are included in the study data comparable to that given for negroes are tabulated separately.

Having assembled the data, Dr Shuey attempts an evaluation of them. The conclusions to be reached from this evaluation are inescapable. Everywhere, and with remarkable consistency, the results indicate a discrepancy between the average intelligence of negroes and of whites in favour of the latter, irrespective of whether the comparison is between actual groups of negroes and whites tested, or between negroes and general population norms. Admittedly the difference is less emphasized amongst the young children, but here it is pointed out that selection factors are involved. It is suggested too that the negroes are often more handicapped by poor environment as they grow older than are whites, while an overall negro deficiency in vocabulary is underlined. If we protest and ask 'But what of the differences between northern and southern states?' we find that this has already been thought of and ample evidence is provided; to take but one example, amongst the Armed Forces in the First World War intelligence ratings fell off in the order northern whites, southern whites, northern negroes, southern negroes. Or again, if we look at Deviates, Dr Shuey estimates that amongst the gifted children there were only one-sixth to one-third as many negroes as whites, whilst amongst retarded children the negroes outnumber the whites by 6 to 1. Social and cultural factors too have been taken into consideration, so the case seems proved.

Now there is no reason why negroes should not differ from whites in intelligence. They differ from whites in pigmentation, in bone structure, in distribution of blood groups, and no doubt in other ways. But where intelligence is concerned we must be careful, not only because of the slur likely to be cast since high intellectual potential is at a premium in the world today, but also because intelligence, since it can only be measured indirectly, is difficult to measure accurately.

To compare negroes and whites impartially where intellectual ability is concerned would require the construction of special batteries of tests for which an item analysis would require to be carried out separately on equal stratified samples of negroes and whites and the items subsequently chosen so as to give both races an equal chance of success. Some of the investigators whose work is reported were aware of this; Young and Bright, for example, in their 1954 study, question the suitability of the W.I.S.C. for southern rural negro children. Although Dr Shuey reports this and draws attention to a few investigators who noted some test-item differences between the racial groups, she does not lay sufficient emphasis anywhere on the possible limitations of the tests employed. Possible bias on the part of testers is another factor which, in a well-planned investigation, would require to be taken into account: a careful examination of the data which Dr Shuey presents suggests that bias may well exist. However, these are points to be kept in mind when evaluating future studies: for the present, we must congratulate Dr Shuey on the interesting review she has given us of the literature to date.

A. E. MAXWELL

The Individual Psychology of Alfred Adler: a Systematic Presentation in Selections from his Writings. Ed. by HEINZ L. ANSBACHER and ROWENA R. ANSBACHER. London: Allen and Unwin. 1958. Pp. xxiii+503. Price 30*s*.

Adler's numerous books consist largely of written-up lectures and republished articles. They are full of valuable material, but they are diffuse and repetitive, and none of them contains a systematic exposition of Individual Psychology. He is thus a particularly suitable subject for the sort of treatment he has received in the book under review. It consists of extracts from his writings, grouped into sections and subsections, interspersed with editorial comment, and prefaced by a short introductory essay. The editors, Professor Heinz and Dr Rowena Ansbacher, have both had extensive personal contact with Adler, and have a strongly sympathetic, though not uncritical, attitude towards Individual Psychology, which they describe as 'a socially oriented, subjectivisitic, holistic psychology of attitudes', as contrasted with Freud's 'biologically oriented, elementaristic, objective drive-psychology'.

The editors have done a remarkable job with scissors and paste. It is not until one has studied the bibliographical note on p. xiv, and mastered the detailed but unobtrusive system of source-referencing, that one realizes that the selections from Adler's writings consist, for the most part, of mosaics of short extracts from different sources, so skilfully put together that the joins hardly show. For example: the chapter on 'Understanding and Treating the Problem Child' begins with a subsection, occupying just over a page, entitled 'The Dynamics of Problem Behaviour'. This contains five paragraphs, drawn from four different sources—two from *Social Interest*, one from *Understanding Human Nature*, and two from previously untranslated articles. Adler, incidentally, has been badly served in the past by his translators, and the present editors have often preferred to translate afresh, even when quoting from books, like the two above-mentioned, of which standard English versions exist.

It would be a pity if anyone interested in Individual Psychology were to confine himself entirely to the book under review, as something of Adler's flavour has evaporated in the rigorous boiling-down process to which he has been subjected. But for sheer nutritive content, no single book by Adler can compare with the present selections. Students of Individual Psychology—and still more, perhaps, those who lecture on it—have good reason to be grateful to the editors.

MARGARET KNIGHT

Function Fluctuation. By C. C. Anderson. *British Journal of Psychology Monograph Supplement No.* 30. London: Cambridge University Press. 1958. Pp. vii + 104. 21*s*.

This monograph takes up an important problem in mental testing on which there is a surprising amount of speculation but relatively little evidence. As is well known, the fact that two tests of one and the same function fail to correlate perfectly is due partly to errors associated with the test, and partly to 'function fluctuation', which is a term originally suggested by Thouless to denote a day-to-day quantitative variation in the function tested. The first half of this monograph reviews the literature on this and related concepts and tries to clarify the undoubted confusion of terms and concepts existing in this field. Also included in the first part is a chapter giving results of previous research. The review is clear and to the point, although at times it appears a little niggling and ungenerous to some of the earlier workers. There is also a superabundance of quotations more characteristic of a Ph.D. thesis than a monograph; these do not always contribute greatly to the argument. However, taken by and large, the reader of this section should have a fairly clear grasp of the problems and theories in this field. Not all relevant material is included; thus, no mention is made of the work of Fleischman, or of the Hartshorn and May studies. The second part of the monograph deals with the author's own experiments which appear to have been unusually extensive as far as the number of tests administered is concerned. Analysis, presentation and discussion are clearcut and incisive. The main results are that function fluctuation is active both in cognitive and in non-cognitive functions, but more extensively in the latter. Function fluctuation does not appear to be a characteristic feature of a given individual; correlations between function fluctuation on one test and function fluctuation on another appear to be nearly zero. From these results the author derives a number of conclusions which are sensible, and a theory of cognitive function fluctuation which does not seem a great advance to the reviewer and which, if true, would suggest that correlations of function fluctuation should be obtained between different tests. However that may be, we are in debt to the author for both his review and his experimental contribution to an important aspect of test psychology that has been somewhat neglected in the past.

H. J. EYSENCK

Electronic Instrumentation for the Behavioral Sciences. By Clinton C. Brown and Rayford T. Saucer. Springfield, Illinois: Charles C. Thomas. Oxford: Blackwell. Pp. xiv + 160. 42*s*.

There is a growing need for a work which achieves what this book promises, a simple account of electronic theory, valves, transistors, circuitry and measuring instruments for research workers who have no previous knowledge of electronics. The need remains, and to fill it demands a degree of competence and clarity not displayed here. A few samples of the book's faults are as follows. Factual errors—R.M.S. value is about 70% of peak to peak value; frequency is number of reversals per unit time: a sine wave has a uniform rate of change. Confusion of symbols—in a single diagram may be found MMF for $\mu\mu$F, MH for mH, MA for mA, M for MΩ, K for kΩ, 8 for 8 μF, 100 for 100 $\mu\mu$F, 6·3 for 6·3 V, 47 for 47 Ω (these interpretations and corrections should not be left to a beginner), elsewhere mf, mfd or MFD for μF. Equations with undefined terms, terms wrongly identified in the text, and no units. Use of advanced concepts with little or no explanation—transients, decibels, phase shift, hard radiation, work function,

thermal noise (who is familiar with these yet needs a simple account of Ohm's law or reactance?). Solecisms—charge equal and opposite to impressed EMF; R_1, R_c, etc. referred to in the text but not marked in the circuit diagram; incorrect explanations of grid leak bias and ohmmeter. There are also contradictions, unclear deductions, false impressions, grammatical errors, misspellings, misprints, faulty diagrams, text references to the wrong figures, and at least one wrong reference in the brief index.

The potential reader can get an accurate, lucid, non-mathematical account of much of the contents, at a third of the price, in *Foundations of Wireless*, by M. G. Scroggie (Iliffe).

J. R. TROTTER

Personality and Motivation Structure and Measurement. By RAYMOND B. CATTELL. London: Harrap. 1958. Pp. xxiv+948. 55*s*.

This is a whale of a book in more ways than one. Its weight and size are impressive; for his 55*s*. the reader gets 948 pages of text, including 661 references of which exactly 100 refer to Cattell's own work. The book weighs 2 lb. 1½ oz., which works out at 1 shilling an ounce, roughly. Again, just as the whale, although living in the sea, is not a fish but a mammal, so Cattell, although writing on personality and motivation, is not at all typical of the modern clinical psychologist. The stress in the title lies on the words 'Structure' and 'Measurement', and for the determination of the former and the accomplishment of the latter Cattell relies very much on complex statistical methods, in particular oblique factor analysis (i.e. a method of factor analysis in which factors need not be independent of each other, but may themselves be correlated) and the discovery of higher-order factors (i.e. factors derived from the intercorrelations of the original oblique factors).

The book is, in a way, a continuation of Cattell's previous volume entitled *Description and Measurement of Personality*; as Cattell says 'For the exact definition of such concepts as *ability*, *interest*, *attitude* and *temperament*...and for examination of evidence on factors for which only final patterns are given in this book, the researcher should consider *Description and Measurement of Personality* as Volume I of the present book'. This dependence on former publications is characteristic of the book throughout; I doubt if readers not familiar with a good deal of Cattell's previous writing could follow the argument at all well. This feature is a little unfortunate, particularly in a book of such length. It will certainly work against the author's hope that it will be useful for undergraduate instruction. Another reason why it may be doubted whether any but the most exceptional undergraduates can be recommended to read this book is related to Cattell's use of factor analytic techniques. While he does attempt in the introductory chapters to introduce these techniques at an elementary level, he soon soars into the highest layers of complexity and argumentation where anyone but the most expert factor analyst would be left gasping for breath. It is not argued that such discussions should not be given, but the attempt to present them in the same book as simple introductions to correlational analysis for the quite uninitiated reader, causes a certain inbalance and an unnecessary lengthening of the book which could have been avoided.

Apart from such relatively trivial matters, what are the reader's first impressions of the book likely to be? Almost certainly he will be impressed by the tremendous enthusiasm and surgency running through the whole research enterprise of which this is a record, and he will admire the quite unusual fertility of ideas which is apparent on almost every page. No one interested in the field could read the discussion beginning on page 436 and culminating in Cattell's 'List of Principles for measuring manifestation of Motivation' without feeling that he is here encountering a writer extremely well read in the psychological literature bearing on personality measurement, and capable of generating novel and original ideas with uncommon facility. Cattell seems almost incapable of touching any subject without changing one's way of looking at it or without throwing up ideas, suggestions and research plans deserving of the most serious consideration. To the expert, therefore, this book will certainly be required reading, and no one working in this field can afford to isolate himself from the invigorating breath of air from Illinois.

There is, however, another side to this medal, and indeed the existence of this other side might have been predicted in terms of Cattell's own factors. The person high in surgency according to Cattell is energetic, original and assertive, but he is also hasty and lacking in self-criticism. The second impression which the reader is likely to carry away from perusing this book is that the ideas generated by Cattell, and the researches carried out by him, are very uneven in value. Some, particularly those relating to personality source-trait structure, are almost certainly valuable additions to knowledge and cannot be neglected by theorists or practical workers in this field. Others, particularly those dealing with dynamic structure, seem to the reviewer of relatively little value or interest. Indeed, on occasion the reader may wonder whether Cattell is quite serious in what he says. He claims, for instance, that drive structures can only be determined by factorial methods, and criticizes the usual work in this field done by experi-

mental psychologists. These criticisms are not entirely without value, but Cattell's own derivations of such concepts as the 'sex erg' do not impress the reviewer as being in any way superior. Essentially the 'sex erg' is defined by a pattern of intercorrelations found among certain attitudes such as 'I want to fall in love with a beautiful woman', 'I want to satisfy my sexual needs'; 'I like to see a good movie now and then'; 'I want to listen to music'; 'I want to travel and explore the world'. The reader must form his own opinion as to whether either method or results can seriously be said to mark an improvement on the work done by learning theorists in isolating and defining drives. One must agree with Cattell that drives can only be defined in terms of some kind of specification equation, but few people, I think, would agree with him that this equation should be a factor analytic one rather than, say, of the type introduced by Hull.

The same feeling of a lack of critical evaluation strikes the reader in considering Cattell's views about the different methods of factor analysis as demonstrated in his 'co-variation chart' on page 494. Correlations between tests, and their analysis by some form of factorial method, are almost certainly useful and valuable adjuncts to more strictly experimental work. Whether correlations between persons or between occasions, or any of the other combinations suggested by Cattell, can be of much value, except in isolated cases, remains to be proven. Many good critics are sceptical, and Cattell's enthusiasm for these methods, interesting as they are on the technological level, hardly seems justified by the results achieved.

Cattell's book, therefore, must be read critically; the reader must never forget that the ideas he meets are thrown off like a shower of sparks by a catherine wheel. There is an insufficiency of insulating material to prevent the poor and mistaken ideas from coming to the surface and thus dominating the good and valuable ones. As with many 'surgent' people, the 'censor' has been nodding. This makes an unbiased review of Cattell's book very difficult and I am sure that many readers will fail to appreciate its outstanding contributions by concentrating their critical interests on the more doubtful ones.

What then is Cattell's main contention? He claims to have isolated some sixteen main factors in the fields of what he calls *L*-data and *Q*-data, by which he means life record variables, such as behaviour ratings, and questionnaire data. He claims that the factors found in one of these fields correspond well with those found in the other on the whole. He also claims the discovery of a number of factors from the intercorrelations of *T*-data, by which he means objective test variables other than questionnaires. These factors do not agree well with those in the *L* and *T* fields. Cattell names his factors, sometimes by using psychiatric terms such as cyclothymia, and sometimes by coining new terms which are liable to cause a headache to most readers. Among these newcomers to the psychological scene are such things as parmia (para-sympathetic immunity), threctia (threat reactivity), prensia, coasthenia, comention and protension, antia, corticalertia and many others. A special set of questionnaires has been published to measure these variables, although unfortunately little seems to be known about their correlations with the more orthodox Guilford factors or many of the other currently used clinical tests.

Most of these factors had already been presented by Cattell previously. He now, for the first time, takes seriously the duty of anyone working with oblique factors to calculate second-order factors, and in doing so he arrived at a conclusion which, to the reviewer, appears a very happy and satisfactory one, namely that 'the publication of extensive analyses...yields two very striking findings—the general integration factor and the extraversion/introversion factor'. Critics have sometimes pointed to the apparent differences in factor structure between Cattell's sixteen factors and the reviewer's set of just two, to wit, neuroticism and extraversion/introversion, as evidence of the subjectivity of factorial methods. In *The Structure of Human Personality* it was argued at great length that extraversion/introversion and neuroticism were conceived of as second-order factors, whereas Cattell's work dealt entirely with oblique first-order factors. It was also suggested then that if the appropriate analysis were done, there was no reason to anticipate any disagreement between the two sets of results. Cattell's finding of extraversion/introversion and integration or adjustment as second-order factors from his own work, suggests that this hypothesis was correct, and it would seem to follow that factorial procedures are perhaps more reliable and valid than some critics have claimed. It may be of interest to note here that, in order to check the identification of Cattell's and his own factors, the reviewer carried out a separate experiment for the purposes of this note, correlating Cattell's measures of these two second-order factors and his own; the resulting correlations were positive, high and highly significant.

Even in an extended review there is not enough space to enter into any further discussion or criticism of Cattell's work. A few points may be worth mentioning, however, in concluding this review. In the first place, the reader will find many examples of Cattell's exceptional verbal facility and his often somewhat biting wit. To quote just one brief example, 'Dexterity in handling a crossbow, or an interest in Guatemalan centipedes, contains specifics little used in High School counselling'. This mordant wit is used with great effect in his discussion of modern clinical, educational and industrial practices; with most

of his criticisms here it would be difficult to disagree, and one can only hope that the recipients of these guided missiles will be in place to receive them.

In the second place, Cattell is essentially a 'one-woman man'. He has remained faithful to factor analysis from his very first writings, and he disdains the attractions of other methodologies and theoretical systems based on them. Thus several times he goes out of his way to criticize modern learning theory and experimental psychology in general in a manner which is not likely to bring about that very essential *rapprochement* between experimental and psychometric methods so much needed and so strongly and well argued for by Cronbach in his recent presidential address to the A.P.A. The reviewer feels certain that Cattell is mistaken in these diatribes and that factor analysis, far from being an adequate method for solving all or most of psychology's problems, is only one method among many with which the psychologist should be familiar.

In the third place, there are certain technical points with regard to which one may feel a certain amount of disquiet. Cattell has always extracted a rather large number of factors from matrices of intercorrelations based on relatively small numbers of subjects. In view of the smallness of the original correlations, one is often doubtful about the significance of any but the first few factors. In his use of criteria for deciding when to stop the extraction of further factors, Cattell has always shown a preference for the least stringent criteria; such a choice may be thought to be mistaken. Cattell criticizes other writers for extracting too few factors, thus giving themselves too little space for rotation; it may be that Cattell extracts too many factors and allows himself too much elbow room in this respect.

The fourth point is related to this. Once having extracted his many factors, Cattell's procedure for identifying factors extracted from different matrices is again rather lenient and may leave the reader unconvinced. It might be interesting to introduce the new methods put forward by Ahmavaara into Cattell's realm of factors to see to what extent his claims are substantiated; unfortunately, his method does not appear to have been used by Cattell.

In his use of experimental methods for the measurement of personality, i.e. in the field of *T*-data, Cattell again shows much ingenuity in adopting techniques of flicker fusion measurement, dark adaptation, and so forth, to the assessment of individual differences. It is often difficult, however, to follow his interpretation of these measurements, because they tend to be of a rather rough and ready kind. There seems to be little control over what may be very important factors, such as pupil size, and the measurement itself is often heavily curtailed in order to make possible the administration of a large number of tests. This curtailment almost certainly introduces uncontrolled variables which may seriously interfere with proper measurement and determination of such variables as dark adaptation.

This leads me to the last of these points. Cattell might reply that as long as the experimental situation is identical for everybody, it *fulfils* the demands of the usual definition of a standard personality test. This, up to a point, is true, but it leaves out of account the equally important question of how results obtained in this manner are to be explained in terms of psychological theory. Cattell seems to be disinclined to take this task seriously; he sometimes appears to think that there is little psychological theory in terms of which such results can be explained. However, the very heterogeneous groups of tests which delineate his *T*-factors, and their failure to show much agreement with his *L*- and *Q*-factors may make him pause and reconsider this question. To find that poor dark vision correlates with neuroticism is interesting, but it is only a beginning; as Thurstone pointed out 'the coefficient of correlation is a confession of ignorance'. It tells us that there is some kind of causal relationship in our data, but it does not tell us the precise nature of this relationship. Is there a physiological inferiority involved? Does the anxiety produced by the testing situation interfere with the solution of the problem set by the test in the manner suggested by Sarason? Is there a direct conditioned fear reaction induced by being in the dark which affects the pupil through its autonomic connexions? Is there a failure of the more neurotic person to keep his eyes fixed to the fixation point? These are only some of the many possible hypotheses, and until an answer is found to the question posed by them, it would be dangerous to use measures of dark vision as personality tests. Conversely, could we but find an answer to these questions, it should be possible to increase the reliability and validity of these measures well beyond their present rather unsatisfactory status. Perhaps it is asking too much of one department to concern itself with these problems in addition to those of the more factorial kind; but my own feeling would be that to go on discovering new factors without having an adequate theoretical basis, firmly grounded in experimental psychology, for the existing one, is a policy of putting the cart before the horse. However, on this point no doubt much could be said from both sides, and it is probably unprofitable to go any more deeply into this question.

In summary then, as I said at the beginning, this is a whale of a book—fascinating, having a rather strange and unusual taste, and on the whole perhaps a little indigestible—nevertheless, something which excites admiration in many of its aspects.

H. J. EYSENCK